

# A többváltozós Shapiro-Wilk tesztek vizsgálata

Pataki Attila  
E-mail: attila.pataki@khb.hu  
(*Ph.D. Értekezés*)

2001. szeptember 30.

# Tartalomjegyzék

<b>Bevezetés</b>	<b>2</b>
<b>1. A normalitás fogalma</b>	<b>9</b>
1.1. Definíciók, tételek . . . . .	9
1.2. A többdimenziós normalitás tesztek . . . . .	19
<b>2. A Shapiro-Wilk teszt, mint a legjobb omnibusz teszt</b>	<b>21</b>
2.1. Az egydimenziós eset . . . . .	24
2.2. A többdimenziós eset . . . . .	26
<b>3. A <math>p</math>-érték kiszámítása az egydimenziós esetre</b>	<b>28</b>
3.1. Royston eljárásai . . . . .	28
3.2. Törtkitevős polinomok . . . . .	29
<b>4. A <math>p</math>-érték kiszámítása a többdimenziós esetre</b>	<b>31</b>
4.1. Az egydimenziós esetben alkalmazott módszer adaptálása . .	33
4.2. A módszer módosítása . . . . .	33
4.3. Egy másik megközelítés . . . . .	36
4.4. Az egyes megközelítések értékelése . . . . .	38
<b>5. A Shapiro-Wilk teszt egyéb többdimenziós kiterjesztései</b>	<b>39</b>
5.1. Rövid áttekintés . . . . .	39
5.2. A próbák erejének vizsgálata . . . . .	42
5.3. A valódi többdimenziós Shapiro-Wilk teszt . . . . .	49
5.4. Következtetések . . . . .	52
<b>6. Alkalmazások</b>	<b>53</b>
6.1. Opcióárazás numerikus módszerekkel . . . . .	54
6.1.1. A Black-Scholes-féle opcióárazás és az „ideális“ piaci feltevések . . . . .	56
6.1.2. Opcióárazás tranzakciós költségek mellett . . . . .	61
6.1.3. Az opcióár becslése variancia csökkentő módszerek hasz- nátával . . . . .	68
6.2. Kockáztatott érték számítás . . . . .	82
6.2.1. Dow Jones részvények . . . . .	91
6.2.2. BUX részvények . . . . .	97
6.2.3. A próbák erejének vizsgálata . . . . .	101
6.3. Következtetések . . . . .	103
<b>I. Függelék</b>	<b>105</b>
<b>Irodalomjegyzék</b>	<b>108</b>

## Bevezetés

A többváltozós eloszlások elmélete — ideértve a normális eloszlást is — a matematikai statisztika egy hallatlanul érdekes területe. Habár elméletük a századfordulóig nyúlik vissza, és a többdimenziós normális eloszláshoz kapcsolódó statisztikai kutatások mögött is legalább 50 év áll, az alkalmazók még a mai napig is vonakodnak a szükséges teszteket elvégezni. A téma választásának is ez a tény volt a legfőbb oka.

Közgazdasági tanulmányaim során mikor sokváltozós statisztikai alkalmazással találkoztam, a többdimenziós normalitás kérdése mindig homályban maradt. Pedig számtalan igen elterjedt sokváltozós módszer, úgy mint a faktor analízis, a főkomponens elemzés, a diszkriminancia analízis, a korrelációs modell, a többdimenziós normalitás hipotézisére épít. Ha a hipotézis nem igaz, akkor az elemzés kimenetele bizonytalan, márpedig a többdimenziós normalitás egy igen erős követelmény.

A disszertáció megírásának kiinduló pontja a *Benedek [1999]* tanulmány volt, mely egyben a legtöbb inspirációt és ötletet is adta. A szerző egy variancia csökkentő módszert használt fel pénzügyi modelljének kiszámításához, mely egy sokváltozós statisztikai összefüggésen alapul. Az opcióár eloszlásának várható érték becsléséhez volt szükség egy „alacsony” varianciájú becslő függvényre. Az opcióárakra vonatkozó mintát ugyanis egy számítógépes simulációs modellből generálta a szerző, melynek számításigénye igen magas. Tudjuk, hogy a várható érték egyszerű mintaátlaggal való becslésének

$$\sigma_Y^2 = \sigma_Y^2/n$$

a varianciája, ahol  $n$  a mintaméretet jelöli. A mintaméret növelésével tehát a becslés varianciája csökkenthető, de nem akármeddig, hiszen számítási korlátaink vannak. Ha viszont találunk egy másik olyan változót, az ún. *kontroll változót*, melynek eloszlása és elméleti paraméterei ismertek, valamint korrelál az opcióárral, akkor a korrelációt és a kontroll változó ismert centrumát illetve ingadozását kihasználva a várható értékre vonatkozóan jobb becslést készíthetünk. A legkézenfekvőbb persze lineáris becslő függvényt konstruálni, melyben a linearitást az opcióár és a kontroll változó együttes normalitására vonatkozó hipotézis implicálja.

A tanulmány számomra egyik legnagyobb érdeme, hogy a hipotézis ellenőrzését a kolléga explicite el is végzi. A forrás, mely a felhasznált többdimenziós teszt alkalmazásához legfőbb útmutatóul szolgált, a *Tew-Wilson [1992]* cikk volt. A szerzők a *Shapiro-Wilk [1965]* által bevezetett  $W$ -teszt egy többdimenziós variánsát használták, amint azt *Malkovich és Affifi* definiálták egy 1973-as cikkükben (*Malkovich-Affifi [1973]*). A sokváltozós alkalmazásoknál tapasztalható hiányosságokat jól illusztrálja, hogy a  $W$ -teszten alapuló többdimenziós változat(ok)nak nincsen sem kereskedelmi forgalomban kapható, sem pedig szabadon hozzáférhető implementációja. Ennek hiányában a teszt elvégzéséhez a szerzők maguk implementálták az

eljárást, és a szignifikancia szintek kiszámításához a Monte-Carlo szimulációval határoztak meg küszöbértékeket. Megjegyezzük, hogy az említett variáns a  $W$ -teszt legkorábbi ismert, többdimenzióra való általánosítása. Tanulmányunk központi elemét ennek a változatnak a vizsgálata képezi, mely úgy tűnik, nem került a kutatói közösség fókuszába.

Mindössze 3 évvel korábban publikált Mardia egy cikket (*Mardia [1970]*), mely egy későbbi többdimenziós normalitás teszt család kiinduló pontja lett. Mardia a ferdeség és csúcsosság többdimenziós kiterjesztésére adott javaslatot, mely többdimenziós normális alapeloszlás esetén meghatározható határeloszlással rendelkezik. Ennek nyomán — a  $W$ -teszttel ellentétben — számos későbbi javaslat született a ferdeség és csúcsosság többdimenzióra való általánosítására illetve az ezekből konstruálható tesztekre, úgy mint *Bera-John [1983]*, *Mardia [1980]* és *Mardia-Foster [1983]*.

Mardia az 1980-as cikkében meg is jegyzi, hogy minden új javasolt többdimenziós tesztet a többdimenziós ferdeségen illetve csúcsosságon alapuló tesztekhez vagy a Shapiro-Wilk teszt Malkovich-Affifi-féle kiterjesztéséhez kell hasonlítani (*Mardia [1980]*). Ennek ellenére 1973 után egyetlen cikket sem találtunk, mely a Malkovich-Affifi variánssal az említés szintjén túl foglalkozott volna. Az alapcikkben találunk Monte-Carlo szimulációs vizsgálatokat a próbák erejére vonatkozóan, de a kor technikai színvonalának megfelelően meglehetősen alacsony replikáció szám ( $r = 500$ ) mellett.<sup>1</sup> Az idő tehát haladt, a teszt pedig elfeledtetett. A cikk ugyanakkor javasolt ferdeségre és csúcsosságra vonatkozó többdimenziós kiterjesztéseket, melyeket viszont meglepő módon aztán más kutatók tanulmányoztak és tovább fejlesztettek (*Baringhaus-Henze [1991]* és *[1992]*).

Ettől függetlenül a későbbiekben történt ugyan néhány kísérlet a Shapiro-Wilk teszt többdimenziós kiterjesztésére (*Royston [1983]*, *Mudholkar et al. [1995]*), de ezek a tanulmányok is mellőzték a korábbi variánsokkal történő összehasonlítást. Ennek nem tudni, mi a valós oka, hiszen ezen cikkek mindegyike hivatkozott az eredeti 1973-as cikkekre. A kutatók számára tehát a dolgozat ismert volt, bár mint említettük, inkább más aspektusból találták fontosnak.

A ferdeségre és csúcsosságra, valamint a  $W$ -tesztre épülő többdimenziós tesztek voltak a legjelentősebb eredmények, melyek érdemben elmozdulást jelentettek a többdimenziós normalitás tesztelés irányába. A ferdeséget és a csúcsosságot egydimenziós esetben előszeretettel alkalmazták, hiszen a normális eloszlást jól jellemző mutatókról van szó. Az ún. Jarque-Bera teszt (*Bera-Jarque [1981]*) például a két mutató kombinációját használja fel idősorok teszteléséhez. A teszt érzékenységet mutat az idősorokban

<sup>1</sup>A replikációk száma a mintavételezések számát jelenti. Monte-Carlo szimulációnál az ismételt mintavételezések során a mintákból számított paraméter becslésekre ugyancsak egy mintát kapunk. A félreértés elkerülése végett utóbbi minta méretére mindig mint replikációk illetve makro-replikációk számára hivatkozunk. A későbbiekben ezek a fogalmak a használat során világossabbá válnak majd.

meglévő esetleges autokorrelációra, ezért ezen az alkalmazási területen a legkedveltebb teszt. Az egydimenziós Shapiro-Wilk teszt is igen elterjedt, amit elsősorban az omnibusz tulajdonságának köszönhet. A teszt statisztika két mutatószám hányadosa, mely normális eloszlás esetén ugyanazt az elméleti paramétert, éspedig a varianciát becsüli. A számláló a szórás négyzetének egy szorzó konstanstól eltekintve a normális eloszlásból származó rendezett statisztikákból számított legjobb lineáris torzítatlan (BLUE) becslése, a nevező pedig az átlagtól való eltérés négyzetösszege, vagyis a tapasztalati variancia  $(n - 1)$ -szerese. Tulajdonságait empirikusan behatóan vizsgálták. Szimmetrikus és koncentrált eloszlásokra és a legtöbb csúcsos eloszlásra  $W$  optimális (Shapiro *et al.* [1968], Pearson *et al.* [1976]), és bizonyos esetekben nagyságrendekkel erősebb, mint bármely más teszt. Ugyanakkor hosszan elnyúló eloszlásokra gyakran alulmarad más alternatívákkal szemben (Filliben [1975]). Bevert gyakorlat, hogy egydimenzióban használt és kedvező tulajdonságokkal rendelkező eljárásokat próbálnak magasabb dimenziókra általánosítani, és így megpróbálni a többdimenziós normalitásra vonatkozóan jó teszteket konstruálni.

Az egydimenziós eloszlás csúcossági és ferdeségi mutatóiból konstruált tesztek esetében, ha az alapeloszlás normális, akkor a teszt statisztikáknak rendszerint van valamilyen meghatározható (többnyire  $\chi^2$ ) eloszlása. Ez alapján pedig lehetőség van a teszt empirikus szignifikancia szintjének, más szóval a  $p$ -értéknek a kiszámítására, mely a tesztelés elvégzéséhez elengedhetetlen. Megjegyezzük, hogy a küszöbértékek illetve a szignifikancia számítások csak nagy minták esetében kellően pontosak. A  $W$ -teszt szignifikancia szintjének kiszámításával az egydimenziós esetben viszont kezdetben problémák voltak, hiszen a teszt statisztikának  $n > 3$  esetén ismeretlen eloszlása van. Ebben hatalmas előrelépést jelentettek a Royston-féle cikkek, melyek a gyakorlatban használható eljárást adtak az empirikus szignifikancia szintek kiszámítására. Ennek a kétségkívül igen nagy jelentőségét nem vitatva megjegyezzük, hogy a számítások polinomiális regressziós közelítésekből származtak. Módszertanilag tehát viszonylag egyszerű eljárásokról van szó, s mégis, a többdimenziós változat(ok)hoz mégsem készültek közelítő eljárások a  $p$ -érték kiszámítására.

A jelen tézistervezetnek az előbbiekben felvázolt hiányok pótlása a célja, vagyis szeretnénk megtudni, hogy:

- Milyenek a  $W$ -teszt tulajdonságai a Malkovich és Afifi által javasolt többdimenziós kiterjesztés alapján?
- Hogyan lehet a többdimenziós esetben a teszt  $p$ -értékének kiszámítására gyakorlatban használható eljárást adni?
- Milyen alkalmazási területen számíthat érdeklődésre?

A teszt tulajdonságainak vizsgálatához az általában szokásos  $H_1$  eloszlásokra vizsgáltuk meg a próba erejét, úgy mint többdimenziós  $\chi^2$ , Stu-

dent  $t$ , Cauchy, kevert normális, lognormális és logisztikus eloszlások. A disszertációnak nem célja a többdimenziós normalitással kapcsolatos szakirodalom áttekintése, sem pedig az utóbbi harminc évben született konkrét teszt javaslatok vizsgálata, mivel az önmagában egy vagy több önálló disszertációt kitenne. Viszont a többdimenziós Shapiro-Wilk tesztek „utáni” kutatás eredményeként áttekintünk néhány fellelhető  $W$ -tesztre épülő változatot, és referenciapontként felhasználtuk az általunk vizsgált variáns empirikus próba erejének vizsgálatában. Mivel azonban ez egy közgazdasági disszertáció, nagyon fontos azt megvizsgálni, hogy közgazdasági adatsorokra jellemző eloszlások esetén milyenek a tesztek tulajdonságai. Ezért *bootstrap* technikával két alkalmazási területen is végeztünk vizsgálatokat a próbák erejére vonatkozóan.

A második kérdésre adandó válaszhoz támpontot az említett Royston-féle cikkek adtak (Royston [1982a], [1982b], [1992], [1995]). A szerző számítógépes implementálásra alkalmas eljárásokat adott a tesztel kapcsolatos számításokhoz, ideértve a  $p$ -érték számításokat is. A Royal Statistical Society által karbantartott Applied Statistics Library (*StatLib*) is az általa fejlesztett algoritmust tartalmazza, de néhány statisztikai szoftverbe illetve algoritmus könyvtárba (pl. *Statistica*, *IMSL*) is ezt a változatot építették be. A *StatLib*-ben található algoritmusokra építve készítettem el saját algoritmusaimat.

Ami a harmadik kérdést illeti, a pénzügy tűnik a teszt szempontjából megfelelő alkalmazási területnek. Ez azért egy érdekes szakterület, mert itt lehet a normalitás problémáját legkevésbé a szőnyeg alá söpörni. A normalitásra épülő modellekkel a legfőbb problémát a kiugró elemek (*outlier*-ek), illetve a vastag farkúság okozta. A pénzügytan szempontjából egyéb elliptikus eloszlásoknak, mint például a Student  $t$ , kevert normális eloszlásoknak, valamint az  $\alpha$ -szimmetrikus eloszlásoknak — mint amilyen a Lévy eloszlás is — fontosabb szerep jut, mint a normális eloszlásnak. A disszertációban vizsgált teszt pedig, úgy tűnik, az ilyen típusú eloszlásokra érzékenyebb mint a vizsgálatban szereplő egyéb változatok.

Gyakran azonban az elliptikusság is túl ideális feltevés pénzügyi idősor vektorok modellezéséhez. A pénzügyi adatsorok normalitásának vizsgálatán túl az is célunk volt, hogy alternatív modelleket bemutassunk, melyek nem építenek az együttes normalitás meglehetősen szigorú alapfeltevésére. Az alkalmazott modellezés ugyanis nem merülhet ki azzal, hogy „A” modell nem jó. Az sem célravezető, bár a közgazdaságban elterjedt szokás, hogy a tesztek alapján az „A” modell alapfeltevésai nem teljesülnek ugyan maradéktalanul, aztán pedig nincsen se „de”, sem pedig „ezért”. A korrekt eljárás szerintünk az, hogy az egyszerűbb és szebb „A” modell helyett van egy „B” alternatíva is, vagyis az együttes normalitás problémája nyilvánvalóan összefügg alternatív modellek keresésével. Reméljük, így azokban az alkalmazott közgazdászokban, akik írásunkat olvassák, némileg csökken a frusztráció amiatt, hogy modelljünkben a kiinduló normalitásra tett fel-

tevések a levegőben lógnak, illetve bátrabban nyúlnak alternatív modellekhez, ha ez adott esetben elméletileg indokolt. Demonstrálni szeretnénk, hogy más megközelítések használata nem jelent leküzdhetlen nehézségeket sem módszertanilag, sem számítástechnikailag.

A munka jellegét tekintve nem támaszkodhattunk kereskedelmi szoftverekre. Egyrészt az általunk végzett vizsgálathoz értelemszerűen nem léteznek implementációk, másrészt a Monte-Carlo szimulációs vizsgálatokhoz automatizált szoftver megoldásra van szükség, mely ugyanakkor kellően hatékony is. A legtöbb alap algoritmust (pl. véletlen számok, lineáris algebrai műveletek stb.) a Press és szerző társai által írott könyvből vettük (*Press et al. [1992]*), mely a tudományos életben a *IMSL* algoritmus könyvtárhoz hasonlóan *de facto* standard. Emellett az algoritmusok ANSI C, FORTRAN 77, PASCAL, valamint a párhuzamos számítógépek lehetőségeit is kihasználó FORTRAN 90 nyelveken megkaphatóak viszonylag elérhető áron. Maga a könyv bizonyos feltételek mellett fejezetenként ingyen is letölthető a *www.nr.com* címről. A StatLib könyvtár, mely statisztikai algoritmusokat tartalmaz FORTRAN 77 nyelven, ingyen letölthető kutatási célokra, a leírásuk pedig az *Applied Statistics* (Royal Statistical Society, Series C) folyóiratban kerül publikálásra.

A disszertáció első fejezete a (többdimenziós) normalitáshoz kapcsolódó, a vizsgálat szempontjából érdekes definíciókat és tételeket ismerteti. Olyan fogalmakat szerettünk volna összegyűjteni, melyek segítenek megérteni a többdimenziós (normális és elliptikus) eloszlások struktúráját, és így a normalitás tesztek mögött rejlő logikát. Nem feltétlen fontos a felsorolt definíciókat és tételeket a disszertáció elején részletesen tanulmányozni, de a dolgozatban alkalmazott módszerek megértéséhez mindenképpen szükség van rájuk, így időnként vissza fogunk rájuk hivatkozni. Ennek a fejezetnek a második alpontjában van egy rövid áttekintés a többdimenziós tesztek készítésénél használt megközelítésekről. Módszertanilag a dolgozatnak ez egy fontos része, mivel az általunk vizsgálat tesztek mindegyike egy-egy megközelítést képvisel.

A második fejezet a Shapiro-Wilk tesztet, illetve annak általam vizsgálandó többdimenziós kiterjesztését ismerteti. A fejezet első alpontjában jórészt a *Shapiro-Wilk [1965]* cikkben publikált eredményeket ismertetjük, a következő alpont pedig a Malkovich-Affifi-féle többdimenziós változatot mutatja be a *Malkovich-Affifi [1973]* dolgozat alapján. A cikk egy Ph.D. disszertáció eredményeinek rövid közlése, melyet sajnos nem sikerült megszerezniünk. Így jórészt saját ismereteinkre támaszkodva próbáljuk meg elmagyarázni a teszt mögöttes logikáját illetve a szerző választásának motívációit, mely a szűkre szabott cikkből nemigen derül ki.

A harmadik fejezet a  $p$ -érték törtkitevős polinomokkal való közelítésének módszerével foglalkozik. Itt egy relatíve új technikáról van szó, mely bár módszertanilag egyszerű, mégsem került be eddig a modellezők tudatába. A lényeg az, hogy polinomiális regressziókban nem egész kitevők használatával

változatosabb függvényformákhoz jutunk, mint egész értékű kitevők használata esetén. Egy polinom illeszkedése akkor javítható, ha bevonjuk a magyarázó változó(k) magasabb fokszámú tagjait is. Ekkor viszont csökken a szabadságfok, és nagy ütemben nő a modell variánsok száma, vagyis hogy mely változók milyen fokszámú hatványait használjuk fel. Tört kitevők használata esetén alacsonyabb fokszám mellett is jó illeszkedést tudunk elérni, és így az optimális modell specifikáció kiválasztása is egyszerűbb feladat. A módszertani áttekintés mellett ismertetjük a StatLib-könyvtár R94-es algoritmusát (Royston [1995]), mely az egydimenziós esetben alkalmazott  $p$ -érték számítás számítógépes implementációja.

A törtkitevős polinomok segítségével a negyedik fejezetben elvégeztük a többdimenziós  $W$ -teszt szignifikancia szintjének becslését a Royston-cikkekből illetve az R94-es algoritmusból kiindulva. Az egydimenziós esethez képest jóval nehezebb feladatunk volt a magas műveletigény és a becslés alapjául szolgáló alapadatok jóval nagyobb mennyisége miatt. Emellett többdimenziós esetben a  $W$ -statisztika eloszlása különböző dimenziókban némileg módosul, vagyis eloszlását nemcsak a mintaméret, hanem a dimenzió szám is befolyásolja.

Miután számítógépes algoritmust készítettünk a  $p$ -érték kiszámítására, egy külön fejezetet szenteltünk a teszt empirikus vizsgálatának. Hogy a történet kerek legyen, egy rövid szakirodalmi áttekintés mellett bevontuk a vizsgálatba a  $W$ -teszt egyéb, általunk fellelt többdimenziós változatait is, így lehetőség van az egyes variánsok összevetésére. Ez azért is érdekes, és szerintünk fontos is, mivel az egyes változatok a 1.2. pontban tárgyalt különböző megközelítések valamelyikéhez tartoznak. Az összehasonlítást mesterséges adatokon végeztük, vagyis ismert eloszláscsaládokból vettük a mintákat. Valós adatokon történő összehasonlítást az alkalmazásokon keresztül mutatunk be a rákövetkező fejezetben.

A hatodik fejezetben két pénzügyi alkalmazást mutatunk be. A 6.1. pontban ismertetett első alkalmazás a disszertáció kiinduló pontját képező Benedek [1999] dolgozathoz tartozik, melyben az opcióár várható értékét próbáljuk becsülni (lineáris) variancia csökkentő módszer használatával. Ismertetjük az ott közölt eredményeket, majd az alkalmazást bővítjük a többdimenziós  $W$ -tesztre vonatkozó összehasonlító vizsgálattal. Bootstrap technikát alkalmazva elemezzük a próbák erejét valós adatsorokon, pontosabban egy valós gazdasági szituációt leíró diszkrét szimulációs modell által generált ismeretlen eloszlású mintákon. Mindezek mellett, ahogy ígértük, bemutatunk egy alternatív megközelítést is, melyhez nincsen szükség a többdimenziós normalitás alapfeltevésére. Ahogy várható, az egyes megközelítések között nincsen minőségi sorrend. A normalitásra épülő lineáris modelltől tudunk analitikusan többet mondani, viszont ennek korlátozottabb az alkalmazhatósága.

A másik alkalmazás (6.2. alpont) egy előkészületben lévő tanulmány része (Benedek et.al.[2001]), melynek magam is társszerzője vagyok. Tőzsdei



részvények hozamainak idősor vektoraira próbálunk többdimenziós modelleket illeszteni. A vizsgálat fő célja, hogy a különböző részvények között megfelelő függőségi, kapcsolatszorossági mértéket találjunk, mely a portfólió kockázatosságát valósághűen jellemzi. A problémát az jelenti, hogy a (lineáris) korrelációs együttható vajon felhasználható-e a kockázatkezelésben és az adatsorokra vonatkozóan az együttes normalitás feltevése tartható-e. A lineáris korreláció alternatívájaként bemutatunk más függőségi struktúrákat, mely a kopula fogalmán keresztül lehetőséget ad rugalmasabb, nem feltétlen szimmetrikus eloszlások konstruálására. Habár ezek a modellek módszertanilag nem túl bonyolultak, mégis hangsúlyozni szeretnénk a kiszámításukkal kapcsolatban felmerülő numerikus problémákat. Csakúgy, mint az előző alkalmazás estében, itt is elvégezzük a többdimenziós Shapiro-Wilk tesztek összehasonlító vizsgálatát a modell alapadatain.

A negyedik és ötödik fejezet teljes egészében önálló eredmény. Ezekben a fejezetekben becslést és algoritmust készítünk a többdimenziós Shapiro-Wilk teszt  $p$ -értékének kiszámítására, illetve összehasonlítva más, a  $W$ -teszten alapuló kiterjesztésekkel megvizsgáljuk a tulajdonságait empirikusan. Az 6.1. alponthban az opcióáras modellen elvégzett további vizsgálat a korreláció indukciós módszer felhasználásával szintén saját eredmény. A 6.2. alponthban ismertetett többszerzős tanulmányban a számítógépes implementáció és a modellek kiszámítása ugyancsak egyéni hozzájárulásunk. Emellett mindkét, a 6.1-6.2. alponthokban a többdimenziós  $W$ -teszteknek modellek alapadatain elvégzett összehasonlító vizsgálata a saját számításainkon alapul.

# 1. A normalitás fogalma

## 1.1. Definíciók, tételek

Ezt az alponatot a definíciók és tételek ismertetésének szenteljük. Néhány definíción keresztül szeretnénk megvilágítani, hogy a többdimenziós normalitás milyen strukturális tulajdonságokkal bír. Ez fontos abból a szempontból, hogy a többdimenziós normalitási teszteket jobban megértsük, valamint fontos az alkalmazások megértése szempontjából is. Didaktikailag szükséges az alábbi fogalmakat itt bevezetni, hisz a későbbiek során folyamatosan használni fogjuk őket. Megértésük azonban sokkal egyszerűbb alkalmazásokon keresztül, így ha az olvasó elsőre nem érti a definíciók tartalmát, javasoljuk hogy később térjen vissza, amikor az adott fogalmat illetve tételt használni fogjuk.

Elsőként definiáljuk magát a normális eloszlást.

**1. Definíció.** Az egydimenziós standard normális eloszlás  $N(0, 1)$  a

$$\phi(t) = e^{-\frac{1}{2}t^2}$$

alakú karakterisztikus függvénnyel jellemezhető.

**2. Definíció.** Az  $m$ -dimenziós standard normális eloszlás természetesen az

$$\psi(\mathbf{t}) = e^{-\frac{1}{2}\sum_{j=1}^m t_j^2}$$

alakú függvény szerint adódik.

Jól ismert tény, hogy a normális eloszlás első- és másodrendű momentumai léteznek, és ezek az eloszlást egyértelműen meg is határozzák. A nem centrális esetben a 1. és 2. definíciók a következőképpen fogalmazhatók meg:

**3. Definíció.** Az  $m$ -dimenziós normális eloszlás karakterisztikus függvénye

$$\psi(\mathbf{t}) = e^{i\mathbf{t}'\boldsymbol{\mu} - \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$$

ahol  $\boldsymbol{\mu}$  az eloszlás várható érték vektora,  $\boldsymbol{\Sigma}$  pedig az eloszlás forma paramétere.

Gyakorlati számításokhoz hasznosabb az eloszlást sűrűségfüggvényével megadni. Mint tudjuk, normális eloszlás esetén ez is létezik, ha a forma paraméter pozitív definit. Ez esetben  $\boldsymbol{\Sigma}$ -át *kovariancia* mátrixnak hívjuk.

**4. Definíció.** Az  $m$ -dimenziós normális eloszlás sűrűségfüggvénye

$$K |\boldsymbol{\Sigma}|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})'\boldsymbol{\Sigma}(\mathbf{x}-\boldsymbol{\mu})} \quad (1)$$

ahol  $K$  alkalmas normalizáló konstans.

**5. Tétel.** A többdimenziós normális eloszlás peremeloszlásai is normális eloszlásúak.

**Bizonyítás.** Bármely alap valószínűségyszámítási könyvben megtalálható, ld. például Anderson [1958]. ■

Ennek megfordítása általában nem igaz. Megértéséhez megadjuk a többdimenziós normalitás egy újabb definícióját. Ehhez először bevezetjük a *kopula* fogalmát. Ismeretes, hogy ha  $F(\cdot)$  egydimenziós eloszlásfüggvény, akkor az  $y = F(x)$  transzformált változó eloszlása egyenletes a  $[0, 1]$  intervallumon. Erre alapozva definiálhatunk egy olyan többdimenziós eloszláscsaládot, melynek peremei a  $[0, 1]$  intervallumon egyenletes eloszlások, így teszőleges egydimenziós eloszlásokat kombinálhatunk egy általunk megválasztott függőségi struktúrával.

**6. Definíció.** *Kopula alatt az  $m$ -dimenziós, egyenletes eloszlású peremekkel rendelkező valószínűségi vektor eloszlásfüggvényét értjük. Más szavakkal a kopula egy olyan  $C : [0, 1]^m \rightarrow [0, 1]$  leképezés, ami rendelkezik az alábbi három tulajdonsággal:*

1.  $C(x_1, \dots, x_m)$  minden komponensében szigorúan monoton
2.  $C(1, \dots, x_j, \dots, 1) = x_j$  minden  $j = 1 \dots m$ -re,  $x_j \in [0, 1]$
3. Tetszőleges  $(a_1, \dots, a_m), (b_1, \dots, b_m) \in [0, 1]^m$  vektorokra, ahol  $a_j \leq b_j$

$$\sum_{j_1=1}^2 \dots \sum_{j_m=1}^2 (-1)^{j_1+\dots+j_m} C(x_{1j_1}, \dots, x_{mj_m}) \geq 0$$

$$x_{k1} = a_k, x_{k2} = b_k, k = 1, \dots, m.$$

A kopula fontos szerepet játszik eloszlások konstruálásában. Jelen diszertáció ugyanis nem kizárólag a normalitás kérdéseivel foglalkozik, de alternatív modelleket is be kíván mutatni. A kopula fogalma pedig alapvető a valószínűségi változók függőségének megértésében, illetve nemszimmetrikus eloszlásokra épülő modellek építésében. A kopula fogalmára építve a normális eloszlást is újra definiálhatjuk, mint egydimenziós normális eloszlások kombinációját, ugyanis

**7. Tétel.** Sklar tétele (Sklar [1996]): Legyen  $H$  egy  $m$ -dimenziós eloszlásfüggvény  $F_1, \dots, F_m$  peremekkel. Ekkor létezik  $m$ -dimenziós kopula, vagyis

$$H(x_1, \dots, x_m) = C(F_1(x_1), \dots, F_m(x_m))$$

Megfordítva, ha  $C$  egy  $m$ -dimenziós kopula és  $F_1, \dots, F_m$  eloszlásfüggvények, akkor a fent megadott  $H$  egy  $m$ -dimenziós eloszlásfüggvény  $F_1, \dots, F_m$  peremekkel.

**8. Következmény.** Ha  $H$  folytonos  $m$ -dimenziós eloszlás  $F_1, \dots, F_m$  peremeloszlásokkal és  $F_1^{-1}, \dots, F_m^{-1}$  kvantilis függvényekkel, akkor a

$$C(u_1, \dots, u_m) = H(F_1^{-1}(u_1), \dots, F_m^{-1}(u_m))$$

kopula egyértelmű.

Ha  $H$  nem folytonos, akkor körültekintően kell eljárni a tétel alkalmazásakor. A tétel alapján megadhatjuk a *Gauss-féle* kopula definícióját, mely egydimenziós normális eloszlású peremekkel többdimenziós normális eloszlást eredményez.

**9. Definíció.** A *Gauss-féle* vagy *normális kopula* alatt a

$$C(\mathbf{x}) \equiv \int_{-\infty}^{\Phi^{-1}(x_1)} \dots \int_{-\infty}^{\Phi^{-1}(x_m)} K |\Sigma|^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\mu)' \Sigma (\mathbf{x}-\mu)} dx_m \dots dx_1$$

függvényt értjük.

A kopula lényege tehát abban áll, hogy az eloszlást felbontjuk peremeloszlásokra illetve az ezeket kombináló kovariancia struktúrára. Ha normális eloszlású peremekre nem-gaussi kopulát teszünk, akkor olyan eloszlásokat tudunk konstruálni, melyek nem normálisak normális eloszlású peremekkel.

Az elmélet általánosabb abban az értelemben is, hogy a kapcsolat szorossága nemcsak a kovarianciával (illetve a később definiálandó lineáris korrelációval) adható meg, hanem ettől általánosabb fogalmakkal is.

**10. Definíció.** Legyen  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2]$  egy  $m$ -dimenziós valószínűségi vektorváltozó. Ekkor  $\mathbf{x}_1$ -nek az  $\mathbf{x}_2$ -re vonatkozó feltételes várható értékét, vagyis az

$$f(\mathbf{x}_2) \equiv E(\mathbf{x}_1 | \mathbf{x}_2)$$

függvényt elsőfajú regressziónak nevezzük.

A normális eloszlás nevezetes tulajdonsága, hogy ez a függvény lineáris.

**11. Tétel.** (Anderson [1958]) Az elsőfajú regresszió normális eloszlás esetén lineáris, vagyis ha  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2] \sim N(\mu, \Sigma)$ , akkor

$$\mathbf{x}_1 \sim N(\mu_1, \Sigma_{11}) \quad \text{valamint} \quad \mathbf{x}_2 \sim N(\mu_2, \Sigma_{22}),$$

$\mathbf{x}_1$ -nek  $\mathbf{x}_2$ -re vonatkozó feltételes eloszlása pedig szintén normális

$$\begin{aligned} \mu_{1|2} &= \mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\mathbf{x}_2 - \mu_2) \\ \Sigma_{11|2} &= \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21} \end{aligned}$$

paraméterekkel.

**Bizonyítás.** Tekintsük tehát az

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

particionálást. A feltételes sűrűségfüggvény meghatározásához képeznünk kell a  $|\Sigma|$  determinánst, valamint a  $\Sigma^{-1}$  inverzet, ami az alábbi:

$$|\Sigma| = |\Sigma_{22}| |\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}|$$

$$\Sigma^{-1} = \begin{bmatrix} \Sigma_{11|2}^{-1} & -\Sigma_{11|2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11|2}^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1}\Sigma_{21}\Sigma_{11|2}^{-1}\Sigma_{12}\Sigma_{22}^{-1} \end{bmatrix}$$

Ezt az (1) képletbe helyettesítve az  $f(\mathbf{x}_1, \mathbf{x}_2) = f(\mathbf{x}_1|\mathbf{x}_2)f(\mathbf{x}_2)$  alakhoz jutunk, melyben az első tényező a feltételes sűrűségfüggvény. Ennek várható értéke, vagyis az elsőfajú regresszió  $\mu_{1|2}$ , ami lineáris. ■

A normalitás az általa implikált linearitás, és az ebből fakadó egyszerű kezelhetőség miatt a regressziós elemzések kedvelt alapfeltevése. Ezen kapcsolat szorosságának leírására definiáljuk a *lineáris korrelációs együtthatót*.

**12. Definíció.** Az  $x_j$  és  $x_k$  valószínűségi változó lineáris korrelációs együtthatója az alábbi formájú:

$$\rho(x_j, x_k) \equiv \frac{\text{cov}(x_j, x_k)}{\sqrt{\text{var}(x_j)\text{var}(x_k)}}$$

Létezik a normális eloszlásnál általánosabb eloszláscsalád is, mely a normális eloszlás bizonyos kedvező tulajdonságaival rendelkezik. Ezek az ún. *körkörös és elliptikus eloszlások*. Vannak esetek ugyanis, mikor a normális eloszlásnak csak bizonyos tulajdonságaira (szimmetria, linearitás) építünk, illetve más okból úgy adódik, hogy nem a normális eloszlású modell a megfelelő hipotézis.

**13. Definíció.** Az  $\mathbf{x}$  valószínűségi vektorváltozó eloszlása akkor körkörös (spherical), ha tetszőleges  $\mathbf{U}$  ortogonális transzformációra  $\mathbf{U}\mathbf{x}$  eloszlása azonos  $x$  eloszlásával, formálisan

$$\mathbf{x} =_d \mathbf{U}\mathbf{x}$$

Az ilyen típusú eloszlások karakterisztikus függvénye igen egyszerű, mivel

**14. Tétel.** (Fang. et.al. [1990]) Egy  $m$ -dimenziós  $\mathbf{x}$  valószínűségi vektor eloszlása pontosan akkor körkörös, ha  $\psi(\mathbf{t})$  karakterisztikus függvénye teljesíti az alábbi két — egymással ekvivalens — feltételek egyikét:

1.  $\psi(\mathbf{t}) = \psi(\mathbf{U}\mathbf{t})$  ahol  $\mathbf{U}$  ortogonális
2. Létezik egy olyan  $\phi(\cdot)$  egyváltozós függvény, hogy  $\psi(\mathbf{t}) = \phi(\mathbf{t}'\mathbf{t})$

**Bizonyítás.** Tetszőleges  $\mathbf{A}$  kvadratikus mátrixra  $\mathbf{A}\mathbf{x}$  karakterisztikus függvénye  $\psi(\mathbf{A}'\mathbf{t})$ , mivel

$$E\left(e^{it'\mathbf{A}\mathbf{x}}\right) = E\left(e^{i(At)'\mathbf{x}}\right) = \psi(\mathbf{A}'\mathbf{t})$$

ezért az első állítás szerint  $\mathbf{x} = {}_d\mathbf{U}\mathbf{x}$ . Ugyanakkor  $\psi(\mathbf{t})$  invariáns az ortogonális transzformációkra, aminek a maximális invariánsa  $\mathbf{t}'\mathbf{t}$ , így  $\psi(\mathbf{t})$  a  $\mathbf{t}'\mathbf{t}$  függvénye kell, hogy legyen, vagyis ebből következik a második állítás. ■

Ha  $\mathbf{x}$  karakterisztikus függvénye  $\phi(\mathbf{t}'\mathbf{t})$  alakú, akkor ezt a továbbiakban így jelöljük:  $\mathbf{x} \sim S_m(\phi)$ ,  $\phi$ -t pedig a körkörös eloszlás *karakterisztikus generátor függvényének* nevezzük. Bizonyítás nélkül közlünk még egy fontos tételt.

Legyen  $\Phi_m = \{\phi(\cdot) | \phi(\mathbf{t}'\mathbf{t}) \text{ } m\text{-dimenziós karakterisztikus függvény}\}$ .

**15. Tétel.**  $\phi \in \Phi_m$  pontosan akkor, ha

$$\phi(x) = \int_0^\infty \Omega_m(xr^2) dF(r)$$

ahol is  $F(\cdot)$  eloszlásfüggvény és

$$\Omega_m(\mathbf{y}'\mathbf{y}) = \int_{S=\{\mathbf{x}|\mathbf{x}'\mathbf{x}=1\}} e^{i\mathbf{y}'\mathbf{x}} dS/S_m$$

$S_m$  az egységgömb felülete,  $\Omega_m(\mathbf{t}'\mathbf{t})$  pedig az  $m$ -dimenziós egyenletes eloszlás karakterisztikus függvénye.

**16. Következmény.** Legyen  $\mathbf{x}$  karakterisztikus függvénye  $\phi(\mathbf{t}'\mathbf{t})$  alakú. Ekkor  $\mathbf{x}$  sztochasztikus ábrázolása

$$\mathbf{x} = {}_d r \mathbf{u}^{(m)}$$

ahol  $r$  és  $\mathbf{u}$  függetlenek,  $r \sim F(x)$

Ha  $\mathbf{x}$ -nek létezik sűrűségfüggvénye is, akkor az előbbi összefüggést  $f(\mathbf{x}) = g(\mathbf{x}'\mathbf{x}) = g(x_1^2 + \dots + x_m^2)$  alakban írhatjuk fel, ahol  $f$  és  $g$  sűrűségfüggvények. Ez azt jelenti, hogy a körkörös eloszlások szintvonalai (hiperfelületei) körök (gömbök). Könnyen bizonyítható, hogy a standard normális eloszlás is ilyen.

**17. Állítás.** Az  $m$ -dimenziós standard normális eloszlás körkörös.

**Bizonyítás.** A standard normális eloszlás karakterisztikus függvénye

$$\psi(\mathbf{t}) = e^{-\frac{1}{2} \sum_{j=1}^m t_j^2}$$

így a 14. tétel alapján eloszlása körkörös, generátor függvénye pedig  $\phi(u) = e^{-\frac{u}{2}}$ . ■

**18. Megjegyzés.** Ebben az esetben az  $r = \sqrt{x_1^2 + \dots + x_m^2}$  sugarú hipergömbön a sűrűséget a  $\chi_m^2$  eloszlásból kell kiszámítani, vagyis az összefüggésben szereplő  $g$  leképezés ennek az eloszlásnak a sűrűségfüggvényét jelöli.

A körkörös eloszlásoknál általánosabb eloszláscsaládot alkotnak az *elliptikus eloszlások*, melyek az  $N(\mu, \Sigma)$  eloszlás általánosításai.

**19. Definíció.** Az  $\mathbf{x}$  valószínűségi vektorváltozó eloszlása akkor elliptikus (elliptically contoured)  $\mu$  és  $\Sigma$  paraméterekkel, ha

$$\mathbf{x} =_d \mu + \mathbf{A}'\mathbf{y} \quad \mathbf{y} \sim S_k(\phi)$$

ahol  $\mu : m \times 1$ ,  $\Sigma : m \times m$ ,  $\mathbf{A} : k \times m$ ,  $\mathbf{A}'\mathbf{A} = \Sigma$ , valamint  $\Sigma$  rangja  $k$ .

**20. Tétel.** Az elliptikus eloszlás peremeloszlásai is elliptikus eloszlásúak.

Értelemszerűen  $N(\mu, \Sigma)$  maga elliptikus.

**21. Állítás.** Az  $m$ -dimenziós normális eloszlás elliptikus.

**Bizonyítás.** Az  $\mathbf{x}$  valószínűségi vektorváltozó eloszlása akkor  $N_m(\mu, \Sigma)$ ,  $\Sigma = \mathbf{A}'\mathbf{A}$ , ha felírható a

$$\mathbf{x} =_d \mu + \mathbf{A}'\mathbf{y} \quad \mathbf{y} \sim N_k(\mathbf{0}, \mathbf{I}_k)$$

alakban. Ekkor  $\mathbf{y} \sim S_k(\phi)$ , ahol is  $\phi = e^{-\frac{u}{2}}$ . ■

**22. Megjegyzés.** Ha  $k = m$ , akkor létezik a sűrűségfüggvény is. A  $k < m$  esetben degenerált eloszlást kapunk, mely a tér egy  $k$ -dimenziós altere felett koncentrálódik.

A  $\Sigma$  jelentése továbbra is ugyanaz marad, a lineáris függőség mértékét illetve az eloszlás formáját határozza meg. Azt is szokás mondani, hogy kovariancia kompatibilis az elliptikus eloszlásokkal. Az elliptikus eloszlások családjában a második legfontosabb eloszlás a Pearson VII típusú eloszlás, másnéven a Student  $t$ -eloszlás.

**23. Definíció.** Az  $\mathbf{x}$  valószínűségi vektorváltozó eloszlása akkor követ  $m$ -változós  $\nu$  szabadságfokú (centrális) Student  $t$ -eloszlást, ha

$$\mathbf{x} =_d \nu^{1/2} \mathbf{y} / s \quad \mathbf{y} \sim N_m(\mathbf{0}, \mathbf{I}_m), s \sim \chi_\nu$$

Jelölése:  $\mathbf{x} \sim Mt_m(\nu, \mathbf{0}, \mathbf{I}_m)$

**24. Megjegyzés.**  $\nu = 1$  esetet Cauchy-eloszlásnak hívjuk.

**25. Állítás.** A centrális Student  $t$ -eloszlás körkörös.

**Bizonyítás.** Tegyük fel  $\mathbf{x}$   $t$ -eloszlású, ekkor

$$\mathbf{x} = \nu^{1/2} \mathbf{y} / s = \nu^{1/2} r \mathbf{u}^{(\tau)} / s = r^* \mathbf{u}^{(\tau)}$$

ahol  $r \sim \chi_\tau$ ,  $s$  és  $\mathbf{u}^{(\tau)}$  függetlenek,  $r^* = \nu^{1/2} r / s$  ( $r^* / \tau \sim F_\nu^\tau$ ), így  $\mathbf{x}$  körkörös.

■

**26. Megjegyzés.** Ez esetben pedig  $r$  mentén a sűrűséget az  $F_m^n$  eloszlásból kell kiszámítani.

**27. Definíció.** Az  $\mathbf{x}$  valószínűségi vektorváltozó akkor követ  $m$ -dimenziós,  $\nu$  szabadságfokú,  $\mu$  és  $\Sigma$  ( $\Sigma = \mathbf{A}'\mathbf{A}$ ) paraméterű (nem centrális) Student  $t$ -eloszlást, ha

$$\mathbf{x} = \mu + \mathbf{A}'\mathbf{y} \quad \mathbf{y} \sim Mt_m(\nu, \mathbf{0}, \mathbf{I}_m)$$

Jelölése:  $\mathbf{x} \sim Mt_m(\nu, \mu, \Sigma)$

**28. Állítás.** Az  $m$ -dimenziós  $\mu$  és  $\Sigma$  paraméterű (nem centrális) Student  $t$ -eloszlás elliptikus.

**29. Megjegyzés.** Míg a normális eloszlást a  $\mu$  és  $\Sigma$  paraméterek egyértelműen meghatározták, addig a  $t$ -eloszlás egy  $(\nu, \mu, \Sigma)$  paraméterhármassal adható meg egyértelműen. Az elliptikus eloszlások kompatibilisek a kovariancia (illetve korreláció) fogalmával, de az eloszlás meghatározásához a normális eloszlást kivéve nem elégségesek.

A többdimenziós  $t$ -eloszlás fontos szerepet játszik különösen pénzügyi alkalmazásokban, illetve olyan alkalmazásokban, melyeknél gyakoriak a kiugró elemek illetve extrém események. Függőségi struktúrája összetettebb, mint a normális eloszlásé, ezért rugalmasabban alkalmazható. Emellett speciális esetként ( $\nu = \infty$ ) tartalmazza a normális eloszlást is, és sok tulajdonságában hasonlít is rá.

A kevert normális eloszlás szintén egy igen fontos eloszlás az elliptikus családon belül.

**30. Definíció.** Legyen  $g(\mathbf{x}; \mathbf{0}, \Sigma)$  az  $\mathbf{x} \sim N_m(\mathbf{0}, \Sigma)$  eloszlású változó sűrűségfüggvénye és legyen  $w(v) \geq 0$  monoton növekvő súlyfüggvény, vagyis

$$\int_0^\infty dw(v) = 1$$



Ekkor kevert normális eloszlás alatt az alábbi sűrűségfüggvénnyel rendelkező eloszlást értjük:

$$f(\mathbf{x}) = \int_0^\infty g\left(\mathbf{x}; \mathbf{0}, \frac{1}{v^2} \mathbf{I}\right) dw(v) = (2\pi)^{-m/2} \int_0^\infty e^{-\frac{1}{2} \mathbf{x}' \mathbf{x} / v^2} dw(v)$$

**31. Állítás.** A kevert normális eloszlás elliptikus.

**Bizonyítás.** Tekintsük az

$$\mathbf{x} = w\mathbf{y}$$

sztochasztikus reprezentációt, ahol  $w$  és  $\mathbf{y} \sim N_m(\mathbf{0}, \Sigma)$  függetlenek,  $w(v) \geq 0$  és eloszlásfüggvény. Ebből már az állítás könnyen levezethető. ■

Ez az eloszláscsalád a normális eloszlás egy másik alternatívája. Speciális esetként ( $w(v) = 1$ ,  $\nu = [0, 1]$ ) szintén tartalmazza a normális eloszlást, így a  $t$ -eloszláshoz hasonlóan a normális eloszlás egy másik irányú kiterjesztése, és ugyancsak elliptikus. Analitikusan is viszonylag egyszerűen kezelhető, ezért igen kedvelt eloszlásfajta.

Habár későbbi elemzéseinkhez nem kapcsolható, mégis szót kell még ejtenünk az ún.  $\alpha$ -szimmetrikus eloszlásokról. A körkörös eloszlások egy másik irányban történő általánosításáról van szó, a matematikai statisztika egy kevésbé kutatott, feltöretlen területe. Ide tartozik a normális eloszlás, a Cauchy-eloszlás, valamint egydimenziós esetben a *Lévy-eloszlás*, melynek a pénzügytanban van gyakorlati jelentősége.

**32. Definíció.** Az  $m$ -dimenziós  $\alpha$ -szimmetrikus eloszlás karakterisztikus függvénye

$$\psi(\mathbf{t}) = e^{-c \sum_{j=1}^m |t_j|^\alpha} \quad 0 < \alpha \leq 2, \quad c > 0$$

alakú.

$c = 1/2$   $\alpha = 2$  esetén a normális eloszlást kapjuk, így tehát

**33. Állítás.** Az  $m$ -dimenziós standard normális eloszlás  $\alpha$ -szimmetrikus.

$m = 1$ ,  $0 < \alpha < 1$  esetet a gyakorlatban Lévy eloszlásként emlegetik.  $m > 1$  esetén  $0 < \alpha \leq 2$  sajnos még elég keveset tudunk az eloszlások létezéséről. Az  $\alpha = 1$ , vagyis 1-szimmetrikus eloszláscsaládot vizsgálták eddig behatóbban, megemlíjtjük, hogy a Cauchy-eloszlás is ide tartozik. Ha létezik is az eloszlás, sűrűségfüggvény akkor is ritkán létezik, ezért az alkalmazók eléggé kerülnek ezt az eloszláscsaládot. A témáról ld. Fang et.al. [1990].

A kovariancia és a (lineáris) korreláció az elliptikus eloszlásokkal konzisztens fogalmak. A definícióban megadott forma paraméter ( $\Sigma$ ) jelentése azonos, a lineáris páronkénti függőséget méri. A legfőbb probléma viszont az vele, hogy a mintából számított tapasztalati korrelációs együtthatóval nem mindig becsülhető, mivel utóbbi az eloszlás momentumaira épül. így ha nem létezik az eloszlás első és/vagy második momentuma, illetve az végtelen, akkor ez a kapcsolatszorosság mérőszám nem számítható (ui. a nevező végtelen). Egy másik gyakorlati probléma a minta kovariancia érzékenysége a kiugró elemekre. így a nem normális eloszlású adatok elemzésekor a becsült lineáris korreláció meglehetősen instabil is lehet, ezért célszerű egyéb, robusztusabb becsléseit használni.

*Lindskog [2000a]* cikkében kimerítő empirikus elemzést ad a lineáris korreláció alternatív becslési lehetőségeiről, és ezek viselkedéséről különböző (elliptikus és nem elliptikus) eloszlások esetén. A legtöbb szempontból sikeres módszereket mi is felhasználtuk egy későbbi fejezetben bemutatásra kerülő alkalmazásban. Ezen mutatók definícióit adjuk meg a következőkben.

**34. Definíció.** *Kendall-féle rangkorreláció (vagy Kendall-féle tau) alatt az*

$$\rho_{\tau}(x_j, x_k) = P \left[ \left( x_j^{(1)} - x_j^{(2)} \right) \left( x_k^{(1)} - x_k^{(2)} \right) > 0 \right] - P \left[ \left( x_j^{(1)} - x_j^{(2)} \right) \left( x_k^{(1)} - x_k^{(2)} \right) < 0 \right]$$

összefüggést értjük.

Mint látható a definícióból, ez egy rangkorreláció. Nem függ az eloszlás momentumaitól, így mindig kiszámítható. Fontos szerepe van az elliptikus eloszlások elméletében, mivel szoros kapcsolatban van a lineáris korrelációval.

**35. Állítás.** *A lineáris korrelációs együttható és a Kendall-féle rangkorreláció között elliptikus eloszlások esetén a következő a kapcsolat:*

$$\rho_{\tau} = \frac{2}{\pi} \arcsin \rho$$

Hasznos lehet tehát a lineáris korreláció mellett a Kendall-féle tau használata is, hiszen mindig kiszámítható, emellett értéke stabilabb, mint a mintából számított korrelációé. Például egy  $\nu = 3$  szabadságfokú  $t$ -eloszlásban gyakoriak a kiugró elemek, amire a minta kovariancia igen érzékeny,  $\nu = 1, 2$  szabadságfokokra pedig ki sem számítható, hiszen végtelen a variancia. Ilyen esetekben célszerűbb a lineáris korrelációt Kendall-féle tau-ból kiszámítani. A  $\rho_{\tau}$  sajnos nem torzítatlan becslése  $\rho$ -nak (*Kendall-Gibbons [1990]*), az empirikus elemzések alapján azonban ez a torzítás igen kicsi.

Egy másik jól vizsgázott mérőszám a lineáris korrelációs mátrix becslésére egy iterációs eljárást ad (*Gnanadesikan-Kettenring [1972]*).

**36. Definíció.** Többdimenziós kiegyensúlyozás (trimming) módszere az alábbi:

1. Rendezzük csökkenő sorrendbe az  $\mathbf{x}_i$  ( $i = 1, \dots, n$ ) mintaelemeket a középponttól ( $\hat{\mu}$ ) számított Euklideszi távolság,  $d = (\mathbf{x}_i - \hat{\mu})' (\mathbf{x}_i - \hat{\mu})$  alapján.
2. Legyen  $l = \lfloor \alpha n \rfloor$  és

$$\hat{\mathbf{C}} = \frac{1}{n-l} \sum_{i=l+1}^n (\mathbf{x}_i - \hat{\mu}) (\mathbf{x}_i - \hat{\mu})'$$

Ha  $n-l > m$ , akkor  $\hat{\mathbf{C}}$  pozitív definit majdnem biztosan.

3. Rendezzük csökkenő sorrendbe az  $\mathbf{x}_i$  mintaelemeket a középponttól ( $\hat{\mu}$ ) számított Mahalanobis távolság,  $d = (\mathbf{x}_i - \hat{\mu})' \hat{\mathbf{C}}^{-1} (\mathbf{x}_i - \hat{\mu})$  alapján.
4. Legyen  $l = \lfloor \beta n \rfloor$  és

$$\hat{\mathbf{C}} = \frac{1}{n-l} \sum_{i=l+1}^n (\mathbf{x}_i - \hat{\mu}) (\mathbf{x}_i - \hat{\mu})'$$

Ha  $n-l > m$ , akkor  $\hat{\mathbf{C}}$  pozitív definit majdnem biztosan.

5. Legyen  $\hat{\mathbf{R}} = \mathbf{D} \hat{\mathbf{C}} \mathbf{D}$ , ahol  $\mathbf{D}$  diagonális mátrix  $d_{jj} = 1/\sqrt{\hat{c}_{jj}}$  diagonálisokkal. Ha  $\hat{\mathbf{R}}$  stabil, akkor az eljárás véget ér, ellenkező esetben ismételjük meg a 3. lépéstől.

Az eljárásban  $\alpha$  és  $\beta$  a felhasználó által beállított paraméterek.

A trimming lényege tehát abban van, hogy a páronkénti lineáris korrelációs együtthatókat nem a változó párok kétdimenziós adatsoraiból, hanem szimultán, az  $m$ -dimenziós mintából számítjuk ki. Ha ugyanis az elliptikus alapeloszlás nem normális eloszlás, akkor a tapasztalati lineáris korreláció instabilitásából fakadóan előfordulhat, hogy az  $\hat{\mathbf{R}}$  becsült korrelációs mátrix közel szinguláris lesz! Mint tudjuk, normális eloszlásnál, ha  $n > m$ , akkor ilyen esemény zérus valószínűséggel fordul elő, és  $\hat{\mathbf{R}}$  rendszerint jól kondicionált.

Fontos szerepe van még egy, a kopula fogalmához kapcsolódó, aszimptotikus kapcsolatszorossági mérőszámnak.

**37. Definíció.** Legyen  $x_j$  és  $x_k$  valószínűségi változók  $F_j$  és  $F_k$  folytonos eloszlásfüggvénnyel. A  $x_j$  és  $x_k$  felső farkokban vagy felső szélekben való függősége (tail dependence) alatt a

$$\lambda_u = \lim_{u \rightarrow 1} P[x_k > F_k(u) | x_j > F_j(u)]$$

együtthatót értjük. Az alsó farkokban való függőség értelemszerűen következik.

Az aszimptotikus függőség a (lineáris) korrelációs együtthatóhoz hasonlóan páronkénti függőséget mér. Mint megismertük az előzőekben, az elliptikus eloszlások között csak a normális eloszlást határozza meg egyértelműen a kovariancia illetve korrelációs mátrix. Más elliptikus eloszlások függőségi struktúrája azonban összetettebb. A  $t$ -kopula esetén például a szabadságfok csökkenésével növekszik az aszimptotikus függőség (Embrechts et al. [1999]). Ez nagyon hasznos lehet, ha például extrém eseményeket szeretnénk modellezni. Jó példa erre a tőzsdei értékpapírok hozama. Ha valami rendkívüli esemény rázza meg ugyanis a piacot, akkor ez rendszerint több értékpapírt is érint. Az extrém események hatására ezért több, egyébként esetleg gyengén korrelált papír elkezd együtt mozogni. Az 6.2. fejezetben bemutatott alkalmazásnál pedig majd találkozni fogunk olyan kopulákkal, melyek ezen felül még rendelkeznek ún. *globális függőségi* paraméterrel, vagyis a páronkénti függőség és az aszimptotikus függőség mellett az  $m$  változó együttmozgását is tudjuk jellemezni. A témáról jó áttekintés található az Artzner et.al.[1999] és Embrechts et.al.[1999] cikkekben.

**38. Megjegyzés.** Bizonyítás nélkül megjegyezzük azt nem meglepő tényt, hogy a *gaussi kopula* (illetve a *többdimenziós normális eloszlás*) *aszimptotikusan független*. Ez elég logikus, hiszen a *lineáris korreláció teljes mértékben jellemzi ennek a kopulának a függőségi struktúráját*.

## 1.2. A többdimenziós normalitás tesztek

Az előbbieken láttuk, hogy a normalitás kiterjesztése többdimenziós esetben sokféle módon megtehető, s ez alapján ismertettük a többdimenziós normális eloszlás néhány karakterisztikáját (pl. linearitás, körköröség, normális peremek,  $\alpha$ -szimmetria stb.). Ennek megfelelően a többdimenziós normalitástól való eltérés többféle irányban lehetséges, ezért többféle technika létezik, melyek eltérő módon érzékenyek az esetleges defektusokra. Legjobb univerzális módszer nem létezik, de egy ilyen módszer vélhetően nem is lenne gazdaságosan alkalmazható. Általában a konkrét alkalmazástól függ, hogy melyek azok a tulajdonságok, melyekre a módszer használatához feltétlen szükség van. Ha például piaci kockázatot szeretnénk becsülni, akkor a hosszan elnyúló, vastag szélek „ellen kell védekeznünk”, míg regressziós becsléseknél az eloszlás elliptikussága az elemzés szempontjából kritikus karakterisztika.

A különböző megközelítéseket alapvetően az alábbi három csoportba lehet besorolni (Gnanadesikan [1977], Andrews et.al. [1972]):

- $T_1$  Peremeloszlások normalitásának vizsgálata
- $T_2$  Együttes normalitás vizsgálata
- $T_3$  Egydimenziós vetületeken alapuló tesztek

Az első csoport egy igen passzív megközelítés. Abból indul ki, hogy habár a peremeloszlások normalitása nem implikálja az együttes normalitást, mégis sok esetben a normalitás hiánya a peremeloszlásokban is tükröződik. Ezt a megközelítést azonban semmiképpen sem szabad alábecsülni. Előfordulnak olyan esetek, mikor az együttes normalitás hiánya egyéb megközelítésen alapuló módszerekkel nehezen detektálható, pedig a peremekben jól tetten érhető. Dolgozatunkban egy ilyen típusú tesztet is fogunk vizsgálni. A 6.2. fejezetben bemutatott alkalmazásban például a részvények hozamai jellemzően csúcsos eloszlások, amik a peremeloszlásokon elvégzett formális tesztek is kétséget kizáróan igazolnak, ugyanakkor az együttes normalitásnak mégis fogjuk jelét látni. A 6.1. fejezetben viszont mutatunk egy olyan pénzügyi alkalmazást, mely ennek a megközelítésnek a hiányosságait fogja jól demonstrálni. Habár a kétdimenziós eloszlás mindkét pereme közelítőleg normális eloszlású, az együttes eloszlásuk mégsem az, így az alkalmazott sokváltozós statisztikai módszer torzított becslést fog eredményezni.

A korrekt teszt elvégzéséhez az előző megközelítésnél mégis érzékenyebb módszerekre van szükség, melyek magukat a többdimenziós karakteristikákat veszik figyelembe. A legegyszerűbb és legkézenfekvőbb lenne az illeszkedés vizsgálatra vonatkozó  $\chi^2$  teszt elvégzése. A módszer hátránya, hogy a teszt ereje függ a képzett osztályközök számától, mely magasabb dimenziókban méginkább problémát okoz. Ide tartoznak különböző transzformációs módszerek, melyek segítségével az  $m$ -dimenziós mintát egyszerűbben kezelhető struktúrákra képezzük le. A transzformáció lehet például maga az eloszlásfüggvény, mely a mintát az egység hiperkockára képi le, ha az normális eloszlásból származik. Ugyanakkor Healy [1968] például megmutatta, hogy a

$$r^2 = (\mathbf{x}_i - \bar{\mathbf{x}}) \mathbf{S}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$$

közelítőleg  $\chi_m^2$  eloszlású. Ebben a dolgozatban olyan megközelítést fogunk bemutatni, mely arra a tulajdonságra épül, hogy a peremeknek az egymásra vonatkozó feltételes eloszlása lineáris. Mivel azonban ez nemcsak a normális eloszlás, hanem az elliptikus eloszlások sajátja is, ezért ennek a megközelítésnek is vannak fogyatékságai.

A harmadik csoport az első megközelítésnek egy intelligensebb változata. A peremeloszlások vizsgálatán alapuló tesztek számítástechnikai szempontból igen gazdaságosak, hiszen a műveletigény a dimenzióknak csak lineáris függvénye. A peremek helyett azonban célravezetőbb lehet más, egydimenziós metszetek vizsgálata. Tudjuk, ha egy többdimenziós eloszlás minden lehetséges metszetét megvizsgálva normális eloszlást kapnánk, akkor ebből már következne az együttes normalitás. Viszont bizonyos kitételezett metszetek vizsgálata is elégséges lehet, ami nem feltétlen a perem. A disszertáció vizsgálatának középpontjában a Shapiro-Wilk teszt Malkovich és Afifi által javasolt változata áll, melyben az eloszlás Roy-féle metszetére alkalmazzuk az egydimenziós tesztet (Roy [1958]).

Megemlítjük még, hogy a tesztek kategorizálhatjuk még mint *paraméteres* és *nemparaméteres* tesztek. Az általunk vizsgált tesztek mindegyike paraméteres. Jelen tézistervezetben nem kívánunk foglalkozni a nemparaméteres tesztekkel.

A következő pontban a legjobb omnibusz tesztként ismert normalitási tesztet, illetve az azt kiszámító algoritmusokat ismertetjük. Ezután egy további pontban részletesen foglalkozunk a teszt szignifikanciájának kiszámításával egydimenziós esetben, majd pedig módszert adunk az eredeti, *Malkovich-Affifi-féle* kiterjesztés empirikus szignifikancia szintjének kiszámítására. Ezen eszközre támaszkodva egy további pontban vizsgálatot végzünk a próba erejére vonatkozóan, és összevetjük néhány, a szakirodalomban fellelhető egyéb variánssal.

## 2. A Shapiro-Wilk teszt, mint a legjobb omnibusz teszt

A klasszikus hipotézis tesztelés úgy működik, hogy definiálunk a statisztikai sokaságból vett véletlen mintára egy becslőfüggvényt, mely egy pontosan meghatározható eloszlást fog követni, ha a kérdéses minta a feltételezett eloszlásból származik. Formálisan felállítunk egy  $H_0$ , ún. *null hipotézist* a feltételezett eloszlásra vonatkozóan, illetve egy  $H_1$  alternatív hipotézist, mely fenáll, ha  $H_0$  nem igaz. A klasszikus, Neyman-Pearson-féle metodológia alapján a mintateret felosztjuk két tartományra. Ha a teszt statisztika az *elutasítási* — vagy más szóval *kritikus* — tartományba esik, a null hipotézist elvetjük. Ha pedig az *elfogadási* tartományba esik, elfogadjuk.

Az elfogadási tartomány rendszerint a  $H_0$  eloszlás értelmezési tartományának azon része, mely felett az eloszlás legnagyobb része, többnyire 90, 95 vagy 99% — a koncentrálódik. A teszt statisztika adott mintából számított értéke ui. nagy valószínűséggel (a felsorolt valószínűségekkel) ide fog esni. Ha ez az esemény nem következik be, "hajlamosak vagyunk azt gondolni", hogy a minta nem a feltételezett eloszlásból származik, ennek a valószínűsége ugyanis igen csekély volt, a példaként említett esetekben 10, 5 illetve 1%. Ekkor elutasítjuk  $H_0$ -t, és elfogadjuk  $H_1$ -et. A kritikus tartomány felett kiszámított valószínűséget *szignifikancia szintnek* nevezzük, és általában  $\alpha$ -val jelöljük.

Mivel döntésünket bizonytalanság mellett hozzuk, ezért természetesen döntésünk hibás is lehet. Például a teszt statisztika értéke a kis valószínűségű elutasítási tartományba esik, pedig a feltételezett eloszlásból származik. Ekkor a null hipotézist — helytelenül — elutasítjuk. Annak az esélye tehát, hogy tévedtünk, pontosan  $\alpha$ -val lesz egyenlő. Ez az ún. *elsőfajú hiba* elkövetésének a valószínűsége.

**39. Definíció.** Az *elsőfajú hiba elkövetésének valószínűségét a teszt mére-*

tének nevezzük. Ezzel ekvivalens elnevezés a szignifikancia szint. Jelölése:  $\alpha$

Az elsőfajú hiba valószínűsége tehát pontosan megadható, ugyanakkor kontrollálható is, hiszen  $\alpha$  értékét az alkalmazó határozza meg. Előfordulhat azonban az is, hogy a teszt statisztika adott mintára számított értéke az elfogadási tartományba esik, és mégis a  $H_1$ -ben megfogalmazott eloszlást követi. Ilyenkor megint csak hibás döntést hozunk, elkövetjük az ún. *másodfajú hibát*. Ennek a hibának a nagysága attól függ, hogy a  $H_0$  illetve  $H_1$  eloszlások mennyire fedik át egymást. Értékének pontos kiszámításához tudnunk kellene  $H_1$  egzakt eloszlását, ami általában nem ismert.

**40. Megjegyzés.** A gyakorlatban a másodfajú hiba nagysága egyszerű  $H_1$  alternatívák esetén empirikusan becsülhető Monte-Carlo szimulációs technikával. Amint a későbbi fejezetekben majd bemutatjuk,  $r$  ismételt mintát generálunk egymástól függetlenül a  $H_1$  eloszlásból, és kiszámítjuk az elutasítási hányadot. Ezt egyből kivonva az empirikusan becsült másodfajú hiba értékét ( $\hat{\beta}$ ) kapjuk.

Ha az elsőfajú hiba elkövetésének valószínűségét csökkenteni akarjuk, akkor a másodfajú hiba valószínűségét növeljük, és fordítva. Ez könnyen belátható, hiszen minél nagyobb az elfogadási tartomány, annál nagyobb a másodfajú hiba, illetve annál kisebb a szignifikancia szint, vagyis az elsőfajú hiba. Mivel egy adott eljárás esetén a kétféle típusú hiba csak egymás rovására mérsékelhető, ezért a hibák csökkentése különböző teszt eljárások közti választás révén valósítható meg. Tehát adott eljárások közül azt fogjuk választani, mely adott  $\alpha$ -szignifikancia esetén a legkisebb másodfajú hibát eredményezi, vagyis nagyobb az ereje.

**41. Definíció.** A próba erején annak valószínűségét értjük, hogy a null hipotézist helyesen elvetjük, vagyis

$$\delta = 1 - \beta$$

ahol  $\beta$  a másodfajú hibát jelöli.

Az előbbiek szerint  $\hat{\delta} = 1 - \hat{\beta}$  összefüggés miatt a 40. megjegyzés szerint kiszámított elutasítási hányaddal viszont a próba erejét becsüljük meg, ezt nevezzük a *próbafüggvény empirikus értékének*.

Ha a null hipotézis egy paramétert vagy paraméter együttest fogalmaz meg, akkor a próba ereje alternatív hipotézisben megfogalmazott paraméter érték(ek)től is függ. Ekkor kritériumként megfogalmazható, hogy az adott teszt egyenletesen a legerősebb, vagyis *UMP* legyen. Ez azt jelenti, hogy az összes szóbajöhető  $H_1$ -re összehasonlítjuk a teszteket a próbák erejét tekintve, és azt választjuk legjobbnak, amelyik mindig erősebbnek bizonyul.

Az UMP kritérium alkalmazása meglehetősen szűk területre korlátozódik, többnyire modell paraméterek becsléseinek hipotézis vizsgálatához használjuk. Ha a null hipotézisben eloszlást fogalmazunk meg, akkor alternatív hipotézisként önkényesen választunk meg eloszlásokat. Itt nehezebb kritériumként megfogalmazni, hogy "minden alternatíva esetén a legerősebb legyen", hiszen a gyakorlatban egyik eljárás az egyik  $H_1$  ellenében bizonyul jobbnak, míg a másik eljárás más  $H_1$  ellenében. Ilyenkor az egyes eljárások közti döntést meghatározhatja az, hogy milyen  $H_1$ -et valószínűsítünk<sup>2</sup>. Ha mégis legjobbat, vagyis univerzális eljárást akarunk választani, akkor választhatunk egy olyan eljárást, ami a vizsgált alternatív eloszlások esetén legtöbbször jobbnak bizonyul, mint a többi. Ha létezik ilyen eljárás, akkor ezt *omnibusz teszt*-nek szoktuk nevezni.

A tesztekkel szemben megfogalmazható még két általános kritérium.

**42. Definíció.** *A teszt akkor torzítatlan, ha*

$$1 - \beta \geq \alpha$$

*vagyis ha a próba ereje nagyobb, mint a választott szignifikancia szint.*

Ez azt jelenti, hogy ha a  $H_0$  eloszlásból veszünk végtelen sok mintát, akkor  $\alpha = 5\%$  esetén a minták legalább 5%-át el kell, hogy utasítsuk. Tetszőleges  $H_1$  esetén ugyanis mindenképpen efelett a szint felett kell maradnunk, és a „legrosszabb” (leghasonlóbb)  $H_1$  maga a  $H_0$ .

**43. Definíció.** *A teszt akkor konzisztens, ha a minta méretét növelve a próba ereje 1-hez tart.*

Világos, hogy minél nagyobb méretű a mintánk, annál jobban karakterizálható  $H_1$ . Elvárjuk a tesztől, hogy biztosabban elválassza a két hipotézist.

A vizsgálatunk középpontjában a Shapiro-Wilk statisztika többdimenziós variánsai állnak. Azért választottuk ezt a tesztet, mivel az eredeti, egydimenziós változat, melyet Shapiro és Wilk [1965] vezetett be biometrikai alkalmazásokban, rendelkezik az előbbieken említett omnibusz tulajdonsággal. Az empirikus vizsgálatok szerint a próba ereje nagy egy sor alternatív  $H_1$  eloszlás esetén (ld. például Shapiro et.al. [1968], Filliben [1975] és Pearson et.al. [1977]). Kézenfekvőnek tűnik a kérdés, hogy vajon lehetséges-e hasonlóan jó tulajdonságú többdimenziós változatot konstruálni.

---

<sup>2</sup>Erre jó példa a Jarque-Bera-féle normalitás teszt, mely idősorok esetében igen jól alkalmazható (Bera-Jarque [1981]).



## 2.1. Az egydimenziós eset

**44. Definíció.** Shapiro és Wilk  $W$ -tesztje az alábbi formula szerint számítható:

$$W = \frac{\left( \sum_{i=1}^n a_i y_i \right)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

ahol  $y_i$   $\{i = 1, 2, \dots, n\}$  a rendezett mintaelemeket jelöli,

$$\mathbf{a} = [a_1, a_2, \dots, a_n] = \mathbf{m} \cdot \mathbf{V}^{-1} \left( (\mathbf{m} \cdot \mathbf{V}^{-1}) (\mathbf{V}^{-1} \mathbf{m}) \right)^{-\frac{1}{2}},$$

$\mathbf{m}$  és  $\mathbf{V}$  pedig  $n$  darab standard normális eloszlású rendezett statisztika várható érték vektora és kovariancia mátrixa.

A teszt arra az összefüggésre épül, hogy a teszt számlálója és nevezője normális eloszlásból vett minták esetén ugyanazt a mennyiséget becsüli.

$\sum_{i=1}^n a_i y_i / \sqrt{(n-1)}$  az  $y_i$  szórásának legjobb lineáris torzítatlan (BLUE) becs-

lése, a  $\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 / (n-1)}$  mennyiség pedig a korrigált tapasztalati szórás.

Az  $\mathbf{a}$  vektor antiszimmetrikus, vagyis  $a_i = -a_{n-i+1}$ , továbbá ha  $n$  páratlan, akkor  $a_{[n/2]+1} = 0$ . Igaz ugyanakkor, hogy  $\sum_{i=1}^n a_i = 0$  és  $\sum_{i=1}^n a_i^2 = 1$ .

Egy normalitás tesztől elvárjuk, hogy skála és origó invariáns legyen. Ez azt jelenti, hogy a teszt elvégzéséhez nem szükséges az eloszlás két paramétere ( $\mu$  és  $\sigma$ ), amit egyébként jellemzően nem ismerünk. Az origó invariancia abból következik, hogy nem közvetlenül az  $y_i$ -vel, hanem mindig az  $(y_i - \bar{y})$  mennyiséggel számolunk. A skála invariancia a teszt logikájából adódik, ugyanis egy  $\hat{s}^2 / \tilde{s}^2$  hányados, ahol  $\hat{s}^2$  és  $\tilde{s}^2$  a variancia két különböző becslése, ezért nem lényeges az alapeloszlás szórását ismernünk. A későbbiekben még hasznát vesszük a következő két tételnek, illetve a bizonyításoknak.

**45. Tétel.** (Shapiro-Wilk [1965])  $W_{\max} = 1$ .

**Bizonyítás.** Mivel  $W$  origó invariáns, feltehetjük, hogy  $\bar{y} = 0$ , így

$$W = \left( \sum a_i y_i \right)^2 / \sum y_i^2$$

Mivel

$$\left( \sum a_i y_i \right)^2 \leq \sum a_i^2 \sum y_i^2 = \sum y_i^2$$

felhasználva, hogy  $\sum a_i^2 = 1$ , ezért  $W$  felső korlátja 1. Ez a korlát az  $y_i = \lambda a_i$  pontban el is éri. ■

**46. Tétel.** (Shapiro-Wilk [1965])  $W_{\min} = na_1^2/(n-1)$ , melyet a

$$\mathbf{y} = \left[ \frac{(n-1)}{na_1}, \frac{-1}{na_1}, \dots, \frac{-1}{na_1} \right]$$

pontban vesz fel.

**Bizonyítás.** Mivel a teszt  $y_i$ -ben nullad fokon homogén, ezért a számláló értékét rögzíthetjük, vagyis legyen  $\sum a_i y_i = 1$ . Vezessük be továbbá a  $\sum y_i = 0$  feltételt. Ekkor a nevező,  $\sum y_i^2$  maximumát keressük az előbbi két feltétel által megadott poliéderen. Mivel a kifejezés konvex, ezért a maximumot az alább felsorolt csúcspontok közt kell keresnünk:

$$\begin{aligned} & \left[ \frac{(n-1)}{na_1}, \frac{-1}{na_1}, \dots, \frac{-1}{na_1} \right] \\ & \left[ \frac{(n-2)}{n(a_1+a_2)}, \frac{(n-2)}{n(a_1+a_2)}, \dots, \frac{-2}{n(a_1+a_2)} \right] \\ & \dots \\ & \left[ \frac{-1}{n(a_1+\dots+a_{n-1})}, \frac{-1}{n(a_1+\dots+a_{n-1})}, \dots, \frac{-(n-1)}{n(a_1+\dots+a_{n-1})} \right] \end{aligned}$$

Könnyen belátható, hogy az optimum a legelső csúcspontban van. ■

Az  $a_i$  koefficiensek kiszámítása azonban korántsem egyszerű. Az  $\mathbf{m}$  és  $\mathbf{V}$  kiszámítható Royston [1982c], valamint a Davies-Stephen [1978] cikkekben közölt algoritmusokkal, illetve a Shea-Scallan [1988] által adott módosításokkal. Mivel  $\mathbf{V}$  szimmetrikus, ezért  $n$  méretű minta esetén  $n + n(n+1)/2$  méretű tárhelyre van szükséges és  $n \times n$  méretű mátrixot kell invertálni. Néhány száz elemnél nagyobb méretű mintáknál  $\mathbf{a}$  direkt kiszámítása nem lehetséges.

Shapiro és Wilk 1965-ös cikkének megjelenésekor a koefficiensek egzakt értéke  $n = 20$ -ig volt ismert (Sarhan-Greenberg [1956]), a szerzők  $n = 50$ -ig a további értékekre becsléseket közölnek. 1982-ben Royston publikált egy eljárást (Royston [1982a]), amivel a teszt kiszámítható volt  $n \leq 2000$  méretű mintákra, ugyanakkor eljárásához egy számítógépes algoritmust is mellékel (Royston [1982b]), mely az  $a_i$  koefficienseket automatikusan kiszámolta. Ez azért érdekes, mert a teszt kiszámításához nem volt szükség többé táblázatok használatára.

Az eljárás aztán a későbbiekben nem bizonyult adekvátnak  $n \geq 50$  méretű mintákra, egymástól függetlenül két tanulmány is született a probléma megoldására. Az egyiket maga Royston írta (Royston [1992]) és természetesen algoritmust is készített (Royston [1995]). Az új eljárás már  $n = 5000$  méretű mintákig volt alkalmazható,  $n = 50$ -ig megfelelően közeli értékeket számított a régihez képest, efelett a mintaméret felett azonban a számított értékek egyre távolodtak egymástól. Mahibbur Rahman és Govindarajulu [1997] ezen tanulmánytól függetlenül korrigálta az 1982-es cikket, és

számítástechnikailag egyszerűen megvalósítható eljárást dolgoztak ki nagy mintákra  $n = 2000$ -ig. Tapasztalataink szerint a két eljárás közel azonos eredményeket produkál.

További problémát jelentett, hogy a statisztika eloszlása analitikusan csak  $n \leq 3$ -ra adható meg. A teszt elvégzéséhez szükséges a küszöbértékek meghatározása, melyeket Monte-Carlo szimuláció segítségével nyerhetünk. Ahhoz azonban, hogy a tesztet táblázatok használata nélkül automatikusan, számítógépes algoritmussal kiszámíthassuk, szintén közelítő eljárásokra van szükség. A Royston [1982b] és [1995] cikkekben publikált algoritmusokban implementáltak ilyen közelítő eljárásokat.

## 2.2. A többdimenziós eset

**47. Lemma.** Egy  $\mathbf{x}$  valószínűségi vektorváltozó pontosan akkor követ  $N(\mu, \Sigma)$  paraméterű többdimenziós normális eloszlást, ha  $y = \mathbf{c} \cdot \mathbf{x}$  normális eloszlású  $N(\mathbf{c} \cdot \mu, \mathbf{c} \cdot \Sigma \mathbf{c})$  paraméterekkel, ahol  $\mathbf{c}$  teszőleges, nem nulla konstans vektor.

Ezen egyszerű állításra alapozva a Shapiro-Wilk teszt könnyen kiterjeszthető a többdimenziós esetre.

**48. Definíció.** (Malkovich-Affi [1973]) A többdimenziós normalitás hipotézisét akkor fogadjuk el, ha

$$\min_{\mathbf{c}} W(\mathbf{c}) \geq k_W \quad (2)$$

ahol  $W(\mathbf{c})$  a 44. definíció szerint számítódik az  $y_i = \mathbf{c} \cdot \mathbf{x}_i$  helyettesítéssel,  $k_W$  pedig alkalmas konstans (kritikus érték).

**49. Megjegyzés.** Ez a változat a talán legkorábbi ismert kiterjesztés. A definíció lényege abban áll, hogy elegendő egyetlen kitüntetett metszetet, mégpedig a legrosszabb esetet megvizsgálunk. Ez tehát az 1.2. alpontban megadott csoportosítás szerint a  $T^2$  családhoz tartozik.

A minimum akkor éretik el, ha

$$\begin{aligned} \mathbf{c} \cdot (\mathbf{x}_1 - \bar{\mathbf{x}}) &= \frac{(n-1)}{na_1} & \mathbf{c} \in \mathbb{R}^m, \quad m \leq n-1 \\ \mathbf{c} \cdot (\mathbf{x}_i - \bar{\mathbf{x}}) &= \frac{-1}{na_1} & i = 2, \dots, n \end{aligned}$$

teljesül. Ez a rendszer túldeterminált, a szerzők ezért az *ols* közelítést alkalmazzák, vagyis a következő négyzetösszeget minimálják:

$$\min_{\mathbf{c}} \left[ \mathbf{c} \cdot (\mathbf{x}_1 - \bar{\mathbf{x}}) - \frac{(n-1)}{na_1} \right]^2 + \sum_{i=2}^n \left[ \mathbf{c} \cdot (\mathbf{x}_i - \bar{\mathbf{x}}) - \frac{-1}{na_1} \right]^2$$

**50. Megjegyzés.** Bizonyos  $\mathbf{x}$ -ek mellett persze megoldható, ezért a teszt minimuma továbbra is  $W_{\min} = na_1^2/(n-1)$ .

A feladat megoldásaként  $\mathbf{c} = \frac{1}{a_1} \mathbf{A}^{-1} (\mathbf{x}_1 - \bar{\mathbf{x}})$  adódik. Mivel  $\mathbf{x}_1$  elvben bármelyik lehet, ezért a tesztet az alábbi algoritmus szerint számolják ki:

1. Legyen  $\mathbf{x}_m$  az a mintaelem, mely a

$$(\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_m - \bar{\mathbf{x}}) = \max_{1 \leq i \leq n} \{(\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})\}$$

maximumot kielégíti, ahol  $\mathbf{A} = \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}}) (\mathbf{x}_i - \bar{\mathbf{x}})'$ .

2. Rendezzük növekvő sorrendbe az  $y_i = (\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$  statisztikákat.
3. A teszt statisztika ekkor a

$$W = \frac{\left( \sum_{i=1}^n a_i y_i \right)^2}{(\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_m - \bar{\mathbf{x}})} \quad (3)$$

szerint számítható, az  $a_i$ oefficiensek továbbra is a 44. definíció alapján számolandók.

**51. Megjegyzés.** Megjegyezzük, hogy az  $a_i$ -k számítása nem függ a dimenziótól. Az eloszlás valamely kitüntetett metszete mentén végezzük el az egydimenziós tesztet, ez az eljárás lényege.

**52. Megjegyzés.** A (3) képletben és a 44. definícióban megadott formulák nevezője valójában azonosság, viszont nyilván célszerű a számítás 3. lépcsőjében amúgy is rendelkezésre álló  $(\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_m - \bar{\mathbf{x}})$  értéket behelyettesíteni. Technikailag tehát csak annyi történik, hogy alkalmas konstans  $\mathbf{c}$  vektorral rendezett statisztikákat számolunk a többdimenziós mintából, amire aztán alkalmazzuk a 44. definíciót.

**53. Megjegyzés.** Ne felejtsük el, hogy a definíció szerint az egydimenziós minimumhoz négyzetes hibában a legközelebbi pontot számítjuk ki, ami nem feltétlen azonos a  $\min_{\mathbf{c}} W(\mathbf{c})$  értékkel. Erre a problémára a későbbiekben még visszatérünk.

A teszt kiszámíthatóságán túlmenően szükség van küszöbértékekre, vagy a szignifikancia szint valamilyen jó közelítésére. Az előző pontban egy mondat erejéig megemlítettük, hogy egydimenziós esetben már van olyan algoritmus, mely számít empirikus szignifikancia szintet. Az viszont korántsem

biztos, hogy többdimenziós esetben is ugyanazok a küszöbértékek jók lesznek, vagyis ezek az algoritmusok nem használhatóak többdimenzióban.<sup>3</sup> A megfelelő percentilis értékek kiszámítása természetesen lehetséges Monte-Carlo szimulációval, de hasznosabb lenne, ha a többdimenziós esetben is tudnánk jól használható algoritmust adni szignifikancia szintek becslésére. A következő pont ilyen becselő eljárások készítésével foglalkozik.

### 3. A $p$ -érték kiszámítása az egydimenziós esetre

Ha a  $p$ -érték kiszámítására nem áll rendelkezésre analitikus megoldás, akkor a legegyszerűbb polinom közelítést alkalmazni. Az első lépés itt is az, hogy Monte-Carlo szimulációval legeneráljuk a teszt adott mintaméretéhez ( $n$ ) illetve adott dimenzióhoz ( $m$ ) tartozó empirikus eloszlását. Ezután pedig regressziós technika alkalmazásával az  $n$  (illetve  $[n, m]$ ) értékének függvényében megpróbálunk becslést adni a megfelelő küszöbértékekre, vagy megpróbáljuk a teszt adott értékéhez tartozó szignifikancia szintet megbecsülni.

Megjegyezzük, hogy ez a terület regressziós módszertanának fejlődését is inspirálta, a vizsgálatok szerint ugyanis jobb illeszkedést kaphatunk, ha az egész kitevős polinomok helyett törtkitevős polinomokat használunk. ígéretes megoldásnak tűnik statisztikai változók kapcsolatának számszerűsítésére.

#### 3.1. Royston eljárásai

A Shapiro-Wilk teszt  $p$ -értékének kiszámítására adott eljárások úgy működnek, hogy a teszt statisztika empirikus eloszlását alkalmas normalizáló transzformációval átviszik közelítőleg normális eloszlásba, majd ezt standardizálva, a kapott  $N(0, 1)$  eloszlásra számítják ki a  $p$ -értéket. Erre ui. léteznek már jó közelítést adó algoritmusok, melyek algoritmus könyvtárakban rendelkezésre állnak.

*Royston ([1982a]) a következő transzformációt használta:*

$$y = (1 - W)^\lambda \quad \text{és} \quad z = \frac{(y - \mu)}{\sigma} \quad (4)$$

ahol  $\lambda$  becsült paraméter,  $\mu$  és  $\sigma$  pedig a normalizált eloszlás várható értéke és szórása. A három paraméter egyike sem ismert, csak becsülhető, mint ahogy magát az eloszlást is szimulációval hozzuk létre. A  $\lambda$  paraméter értékét optimalizálásból tudjuk meghatározni. A  $\mu$  és  $\sigma$  paramétereket ugyan csak becsülni kell, hiszen a normalizáló transzformáció közelítőleg normális eloszlásba viszi át a  $W$  eloszlását, de nem tudjuk, hogy az eloszlásnak mik a paraméterei. Ezt pedig tudnunk kellene, hiszen nekünk standard normális eloszlásra van szükségünk. Technikailag két dolgot kell tennünk. Először is meg kell határoznunk minden  $n$  mintaméretre a megfelelő  $\lambda$ -át, ekkor  $\mu$

<sup>3</sup>Könnyen ellenőrizhető, hogy többdimenzióban mások a küszöbértékek.

és  $\sigma$  már kiszámítható.<sup>4</sup> Másrészt a  $\lambda$ ,  $\mu$ ,  $\sigma$  paraméterek és a mintaméret között kellene a függvényt kitalálni, mivel az algoritmusnak mindhárom paramétert pusztán a mintaméret, vagyis  $n$  ismeretében kell reprodukálnia.

A  $\lambda$  értéke a következőképpen határozható meg. Vegyük a standard normális eloszlást, valamint a Shapiro-Wilk eloszlás néhány kitüntetett percentilisét, mégpedig 0.1, 0.5, 1, 2, 3, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 99%-os értékeket, és számoljuk ki közöttük a korrelációt  $\lambda$  függvényében. Az az optimális  $\lambda$ , ahol a korreláció maximális.

A  $\lambda$ ,  $\mu$ ,  $\sigma$  paraméterek és  $n$  közötti kapcsolatra célszerű polinomiális regressziót becsülni, vagyis egy-egy regresszióval az  $n$  aktuális értékéből rendre kiszámíthatjuk a  $\hat{\lambda}$ ,  $\hat{\mu}$  és  $\hat{\sigma}$  becsléseket. Ekkor tehát  $n$ -ből meg tudjuk becsülni a paramétereket, a paraméterek segítségével pedig a (4) alapján a  $W$  statisztikából készíthetünk egy  $z$  statisztikát.

Ez volt az az eljárás, melyet az AS181 algoritmus használt. Ezt az algoritmust aztán később szerzője korrigálta, mivel az  $a_i$  koefficiensekre adott közelítés nem bizonyult megfelelőnek  $n \geq 50$  esetén (Royston[1995]). Az új, R94-es algoritmusban az egyszerű  $\ln(1 - W)$  transzformációt használta fel, melyhez tehát nincs szükség a  $\lambda$  paraméter becslésére.<sup>5</sup> Ezzel a transzformációval  $n = 12 - 5000$  méretű mintáknál megfelelően lehetett a normális eloszlást közelíteni. Az algoritmus továbbá már nemcsak  $p$ -értéket tudott számolni, de cenzorált mintára is működött.

Megjegyezzük, hogy a fentiekől függetlenül megjelent egy másik tanulmány is (Mahibbur-Govindarajulu ([1997]), mely az  $a_i$  koefficiensek közelítésére eljárást ad. A szerzők azonban nem adnak módszert a szignifikancia szintek kiszámítására.

### 3.2. Törtekitevős polinomok

Royston és Altman [1994] vezette be a polinomok egy kibővített családját, melyet *törtekitevős polinom*-nak neveztek el. A hagyományos, pozitív egész kitevős polinomokkal szemben a kitevők  $\mathcal{P}$  halmazát kiterjesztették negatív és nemegész valós elemekkel is. A törtekitevős polinomok családjában meglévő függvényformák sokkal változatosabbak, mint a hagyományos polinomok családjában. Empirikus eredményeik pedig azt mutatják, hogy egy  $k$ -ad fokú törtekitevős polinom illeszkedése jobb, mint egy azonos, sőt bizonyos esetekben egy magasabb foksámú hagyományos polinomé.

**54. Definíció.** (Royston-Altman[1994]) A  $k$ -ad fokú, egyváltozós törtekitevős polinom alatt az alábbiértjük:

<sup>4</sup>Egészen pontosan csak becsülhető, de a Monte-Carlo szimuláció mérete általában igen nagy. Az így kapott  $\mu$  és  $\sigma$  értékeket a valós paramétereknek tekintjük.

<sup>5</sup>Az  $(1 - W)^\lambda$  transzformáció  $\lambda \rightarrow 0$  határesetre az  $\ln(1 - W)$  transzformáció.

$$\phi_k(X; \xi, \mathbf{p}) = \xi_0 + \sum_{j=1}^k \xi_j X^{(p_j)},$$

ahol  $k$  pozitív egész,  $\mathbf{p} = (p_1, p_2, \dots, p_k)$  valós értékű kitevők vektora ( $p_1 < p_2 < \dots < p_k$ ),  $\xi = (\xi_1, \xi_2, \dots, \xi_k)$  valós értékű együtthatók, valamint a kitevőben a zárójel a Box-Tidwell transzformációt jelöli.

**55. Definíció.** Box-Tidwell transzformáció a következő:

$$X^{(p_j)} = \begin{cases} X^{p_j}, & \text{ha } p_j \neq 0 \\ \ln X, & \text{ha } p_j = 0 \end{cases}.$$

**56. Megjegyzés.**  $p_j = j$ ,  $j = 1, 2, \dots, k$  esetén a  $k$ -ad fokú egész értékű polinomot kapjuk vissza (feltéve persze, hogy  $\xi_k \neq 0$ ).

Az 54. definíció kiterjeszthető oly módon, hogy megengedünk ismétlődő kitevőket is. A kiterjesztés úgy adódik, hogy habár  $k = 2$ ,  $p_1 = p_2$  esetén a

$$\phi_2(X; \xi, \mathbf{p}) = \xi_0 + (\xi_1 + \xi_2) X^{(p_1)}$$

polinom nem másod-, hanem csak elsőfokú, a

$$\lim_{p_2 \rightarrow p_1} \xi_0 + \xi_1 X^{(p_1)} + \xi_2 X^{(p_1)} \frac{X^{(p_2 - p_1)} - 1}{p_2 - p_1} = \xi_0 + \xi_1 X^{(p_1)} + \xi_2 X^{(p_1)} \ln X$$

határérték azonban nem, s így a három paraméteres görbék egy családját kapjuk. A fenti formához a standard  $\phi_2(X; \xi^*, \mathbf{p}) = \xi_0^* + \xi_1^* X^{(p_1)} + \xi_2^* X^{(p_2)}$  modellből a  $\xi_0 = \xi_0^*$ ,  $\xi_1 = \xi_1^* + \xi_2^*$ ,  $\xi_2 = (p_2 - p_1) \xi_2^*$  helyettesítésekkel jutunk el. Az  $k > 2$ ,  $p_1 = p_2 = \dots = p_k$  esetben a fenti határérték  $\xi_0 + \xi_1 X^{(p_1)} + \sum_{j=2}^k \xi_j X^{(p_1)} (\ln X)^{j-1}$ . így az 54. definíció általánosított változata az alábbi:

**57. Definíció.** (Royston-Altman[1994]) Tetszőleges  $p_1 \leq p_2 \leq \dots \leq p_k$  valós kitevők esetén legyen  $H_0(X) = 1$  és  $p_0 = 0$ , valamint

$$H_j(X) = \begin{cases} X^{(p_j)}, & \text{ha } p_j \neq p_{j-1} \\ H_{j-1}(X) \ln X, & \text{ha } p_j = p_{j-1} \end{cases} \quad j = 1, 2, \dots, k$$

Ekkor

$$\phi_k(X; \xi, \mathbf{p}) = \xi_0 + \sum_{j=1}^k \xi_j H_j(X)$$

Több  $X_i$  ( $i = 1, 2, \dots, s$ ) változó esetére a definíció értelemszerűen adódik.

A kitevők halmazára a szerzők  $\mathcal{P} = \{-2, -1, -0.5, 0, 0.5, 1, 2, \dots, \max(3, k)\}$  halmazt ajánlják, mely a gyakorlatban felmerülő problémákhoz elégségesnek bizonyult (Royston-Altman[1997]).

A modellek maximum likelihood módszerrel (ML) becsülhetők. Tekintve, hogy a kitevők csak egy véges  $\mathcal{P}$  halmazán keressük az ML megoldást, a becsléshez használhatjuk a *legkisebb négyzetek módszerét (OLS)*, mégpedig a következők miatt. Ha  $\mathbf{p}$  folytonos változó lenne, akkor a modell nemlineáris lenne a becsülendő  $\mathbf{p}$  és  $\xi$  paraméterekkel, egy előre rögzített  $\tilde{\mathbf{p}}$  esetén azonban a modell már lineáris. Ez esetben pedig (adott rögzített  $\tilde{\mathbf{p}}$ -re) a modell likelihood becslése maga az *OLS* becslés.

A modellválasztás a következőképpen történik. Ha a  $\mathbf{p}$  folytonos változó lenne, akkor a  $\mathbf{p}$  és  $\xi$  paraméterekkel rendelkező modellre kapnánk mondjuk egy  $\hat{\mathbf{p}}$  (*maximum likelihood*) becslést. Jelöljük  $D(k, \mathbf{p})$ -vel a  $k$  fokszámú  $\mathbf{p}$  kitevőjű modell likelihood értékét. Ekkor  $D(k, \mathbf{p}) - D(k, \hat{\mathbf{p}})$  aszimptotikusan  $\chi_k^2$  eloszlást követ, ahol  $\mathbf{p}$  egy tetszőleges (nem feltétlen ML) becslés. Mivel  $D(k, \tilde{\mathbf{p}}) \geq D(k, \hat{\mathbf{p}})$ , ezért  $G = D(k, \mathbf{p}) - D(k, \tilde{\mathbf{p}})$  egy aszimptotikusan használható teszt a  $\mathbf{p}$  megfelelő értékének megválasztásához. A kiinduló modell legyen a lineáris modell  $D(1, 1)$  likelihood-dal, és ezzel állítsuk szembe  $\phi_1(X; \xi, \tilde{p})$  modellt, ahol a  $\tilde{p}$  a  $\mathcal{P}$  halmaz egy általunk tetszőlegesen kiválasztott eleme, ennek likelihood értéke pedig  $D(1, \tilde{p})$ . Ekkor ha  $G = D(1, 1) - D(1, \tilde{p}) > \chi_{1(1-\alpha)}^2$ , az új modellt preferáljuk az előző – ez esetben lineáris – modellel szemben. A becslést és a tesztet aztán megismételjük minden lehetséges  $\tilde{p}$  értékre. A zárójelben elhelyezett  $\alpha$  a szignifikancia szintet jelöli, melynek értékét a szerzők  $\alpha = 10\%$ -nak javasolják megválasztani. A  $k$  és  $k+1$  fokú modellek közti választás esetén pedig a  $\chi_{2(1-\alpha)}^2$  küszöbértéket kell használni.

Többváltozós esetben a fentiekben leírt eljárást először alkalmazzuk az első változóra, vagyis  $X_1$ -re, miközben az  $X_2, X_3, \dots, X_s$  változókat lineárisan szerepeltetjük. A következő lépésben aztán az így kapott legjobb modell specifikációt tekintjük kiinduló modellnek, és  $X_2$  függvényében keressük a legjobb specifikációt most  $X_1, X_3, \dots, X_s$  rögzített specifikációja mellett. Mikor megtaláltuk  $X_s$  függvényében is a legjobb modellt, akkor folytatjuk tovább az  $X_1$  változóval. Az iteráció akkor ér véget, amikor a modell már nem változik.

#### 4. A $p$ -érték kiszámítása a többdimenziós esetre

A  $p$ -érték kiszámítására  $m$ -dimenziós esetben – ismereteink szerint – még nem vállalkozott senki. Ez jórészt abból fakad, hogy ezt a variánst nem vizsgálták és ritkán alkalmazták, részben pedig abból, hogy a teszt statisztika kiértékelése többdimenziós esetben rendkívül számításigényes. Mi ezt a hiányt szeretnénk pótolni, s e célból létrehoztunk egy nagyméretű



adatbázist, melyben a többdimenziós Shapiro-Wilk teszt empirikus eloszlásait tároljuk. A technika elvileg rendelkezésre áll, szükség van egy alkalmas normalizáló transzformációra, valamint megfelelő (törtkitevős) polinom illesztésére. Az itt közölt eredményekből egy cikk is készült (*Pataki [2001]*).

Az eloszlásokat  $r = 10.000$  ismételt mintavételezéssel nyertük. A kiinduló egyenletes eloszlású véletlen számokhoz a *ran1()* eljárást használtuk fel, melynek leírása a *Press et.al.[1992]* műben található. A Park és Miller-féle módszert implementálja a Bays-Durham által adott módosításokkal. Az algoritmus az idézett műben szintén publikált diagnosztikai algoritmusok alapján a legpontosabbnak bizonyult, s a szerzők állítása szerint az összes ismert teszten átmegy. A ciklus ideje nagyjából  $10^8$  körüli, ami az igényeinkhez elégségesnek bizonyul. Az algoritmusnak létezik egy másik, L'Ecuyer által javasolt variánsa is, melyben két pszeudo-véletlen számsorozat összekapcsolásával nagyobb ciklusidejű véletlenszám-generátorhoz juthatunk. Ez az idézett műben mint *ran2()* eljárás szerepel. A diagnosztikai tesztek alapján hasonló pontosságú, mint a *ran1()* eljárás, de ciklusideje  $\sim 2.3 \times 10^{18}$ . Ez a legtöbb gyakorlati problémához elégséges kell, hogy legyen. Adatbázisunk létrehozásához mindkettőt felhasználtuk.

Az egydimenziós normális eloszlásba való transzformációhoz az elterjedt és igen jó tulajdonságokkal bíró *gasdev* algoritmust használtuk, mely a *Bor-Müller* transzformáció révén képez egyenletes eloszlású véletlen számokból normális eloszlású véletlen számokat. Bővebb leírása a *Knuth [1981]* műben, implementációja pedig a *Press et.al.[1992]* könyvben megtalálható. Transzformációs, elutasításos módszer, mely egy normális eloszlású véletlen számhoz átlagosan 2.4 egyenletes eloszlású véletlen számot használ fel.

Ahhoz, hogy független standard normális eloszlású változókból többdimenziós eloszlást kapjunk adott  $\mu$  és  $\Sigma$  paraméterekkel, ahhoz a jól ismert  $\mathbf{x} = \mu + \mathbf{T}\mathbf{y}$  transzformációt célszerű használni, ahol is  $\mathbf{x} = [y_1, \dots, y_m]$  független, standard normális eloszlásból generált változók, és  $\Sigma = \mathbf{T}^T \mathbf{T}$ . Mivel a  $W$ -teszt origó invariáns, ezért a sebesség növelése érdekében  $\Sigma = \mathbf{I}$ -vel számoltunk. Ez az jelenti, hogy ha  $y_j$   $j = 1, \dots, m$  változókat egymástól függetlenül generáljuk normális eloszlásból, akkor  $\mathbf{y} = [y_1, \dots, y_m]$  többdimenziós normális eloszlású  $\mu = \mathbf{0}$  és  $\Sigma = \mathbf{I}$  paraméterekkel.

A számítások dupla precizitással készültek ALPHA AXP 2100-as munkaállomáson. A programot ANSI C-ben implementáltuk, a számítások időigénye így két-három hét volt. A futtatások teljes ideje alatt numerikus hiba illetve programhiba nem állt elő, az adatbázisban sem találtunk instabil értékeket. Az eredményeket — a pontosság megőrzése érdekében — bináris formában tároltuk megőrizve ezzel a teljes gépi pontosságot.

Az adatbázist az R94 algoritmusnak (*Royston [1995]*), illetve a Rahman-Govindaraluju-féle algoritmusnak a Malkovich-Affifi által javasolt kiterjesztésével számoltuk az alábbi esetekre:

$m = 1 - 15$  dimenziókra, valamint  $n = 3 - 100(1)$ ,  $n = 110 - 200(10)$ ,  $n = 250 - 1000(50)$ ,  $n = 1500 - 5000(500)$  méretű mintákra, ahol a zárójelbe

tett érték a lépésközt jelöli.

A két kiinduló egyenletes eloszlású véletlenszám-generátorral és két különböző Shapiro-Wilk-féle implementációval így négyszer generáltuk le az eloszlásokat, az így kapott adatbázis mérete  $\sim 4 \times 120\text{MB}$ .

#### 4.1. Az egydimenziós esetben alkalmazott módszer adaptálása

Az egyes  $(n, m)$  párokhoz tartozó empirikus eloszlásokat normalizáltuk, majd kiszámítottuk az átlagok és szórások adatsorait, ezekre kell majd regressziót felírunk. Az egydimenziós esethez képest most két magyarázó változónk van, éspedig a dimenzió  $(m)$ , mely a regresszióban mint  $X_1$  magyarázó változó szerepel, valamint a mintaméret  $(n)$ , mely az  $X_2$  magyarázó változó lesz. Az egydimenziós  $W$ -statisztika  $p$ -értékének kiszámításánál használt  $y = \ln(1 - W)$  normalizálást használtuk fel a Shapiro-Wilk eloszlás normalizálásához. A transzformáció  $m = 9$ -ig kielégítőnek bizonyult, efelett azonban már nem. A hibák alakulásában a dimenziószám növekedésével szisztematikus változás figyelhető meg, ami azt sugallja, hogy többdimenziós esetben a módszer módosítása szükséges.

#### 4.2. A módszer módosítása

Az első próbálkozás nem eredményezett használható becsléseket, így az empirikus eloszlásokat alaposabb vizsgálatnak vetettük alá. Véleményünk szerint a problémát az okozza, hogy  $m > 1$  esetben a statisztika nem éri el az egydimenziós esetben kiszámított maximumot, az 1-et. Az  $\ln(1 - W)$  normalizáló transzformáció elvileg  $\ln(W_{\max} - W)$ , a Shapiro-Wilk eloszlás egy lognormális eloszlás közelítőleg tükörképe. Többdimenziós esetben is értelemeszerűen  $W_{\max}$ -ot kellene az origóba vinni, feltéve, hogy a közelítő lognormalitás itt is érvényes.

A javasolt módosítás már akkor is meglepően jól működik, ha a maximumot a  $W_{10000} + \varepsilon$  értékkel közelítjük, vagyis az empirikus eloszlások legnagyobb elemével. Azt tapasztaltuk, hogy a statisztikus szempontjából érdekes percentilisekre (0 – 30%) megfelelő az illeszkedés, az  $m$  és  $n$  paraméterek változásával a pontatlanságban nem figyelhető meg szisztematikuság. Az eloszlások közepén válik az eltérés  $\pm 1\%$ -ot elérő nagyságrendűvé, ezek már nem elfogadhatóak.

A tényleges  $W_{\max}$  kiszámítása azonban korántsem egyszerű. Analitikus megoldásra ezideig számos kísérletet tettünk, egyelőre eredménytelenül. Numerikusan elvben lehetséges, a statisztika alkalmas módosításával ugyanis elérhető, hogy a  $W$  legalább folytonos legyen, így  $GO^6$  módszerekkel a maximum egzakt módon kiszámítható.

---

<sup>6</sup> $GO$  néven a vonatkozó szakirodalomban globális optimalizációt értenek. A témáról jó áttekintést nyújtanak a *Pintér [1996]* és *Zhigljavsky [1991]* művek.

Tekintsük a következő optimalizálási feladatot,

$$\max_{\mathbf{x}_1, \dots, \mathbf{x}_n} \min_{\mathbf{c}} W(\mathbf{c}, \mathbf{x}_1, \dots, \mathbf{x}_n) \quad (5)$$

ahol is  $W$  a (2) és (3) képletek szerint adott, azzal a kevés módosítással, hogy az  $y_i = (\mathbf{x}_m - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$  statisztikák rendezése nem szükséges. Könnyen belátható ugyanis, hogy az optimális megoldást olyan minta adja, ahol  $y_i$ -k már sorban vannak. A  $\mathbf{c}$  paraméterben történő minimalizálás is kiváltható a  $(\mathbf{x}_1 - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_1 - \bar{\mathbf{x}}) \geq (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$   $i = 1, \dots, n$  korlátok bevezetésével, így egy folytonos maximum problémával állunk szemben.

Ez a direkt megközelítés a feladat méretéből adódóan a gyakorlatban nem kivitelezhető, mivel a lehetséges megoldások halmaza  $n \times m$  dimenziós, mely több változó esetén nagy mintáknál igen nagy lehet.<sup>7</sup> Ezért egy másik utat kell követnünk.

Vezessük be a  $\mathbf{z}_i = \mathbf{A}^{-1/2} (\mathbf{x}_i - \bar{\mathbf{x}})$   $i = 1, \dots, n$  standardizált változókat. Ekkor tehát  $y_i = \mathbf{z}_i' \mathbf{z}_i$ ,  $\sum_{i=1}^n \mathbf{z}_i = \mathbf{0}$ , valamint  $\sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i' = \mathbf{I}$ . Mátrix algebrai jelölésekkel  $\mathbf{y} = \mathbf{Z}^T \mathbf{Z} \mathbf{e}_1$ ,  $\mathbf{1}^T \mathbf{Z} = \mathbf{0}$ , és  $\mathbf{Z} \mathbf{Z}^T = \mathbf{I}$ , ahol  $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_n]$ .

Most vezessük be a  $\mathbf{B} = \mathbf{Z}^T$  jelölést. A felírás így  $\mathbf{y} = \mathbf{B} \mathbf{B}^T \mathbf{e}_1$ ,

$$\mathbf{B}^T \mathbf{1} = \mathbf{0} \quad (6)$$

$$\mathbf{B}^T \mathbf{B} = \mathbf{I} \quad (7)$$

$\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_m]$ , Shapiro és Wilk tesztje pedig a

$$\cos^2 \varphi = \left\{ \frac{\mathbf{a}^T (\mathbf{B} \mathbf{B}^T \mathbf{e}_1)}{\|\mathbf{a}\| \|\mathbf{B} \mathbf{B}^T \mathbf{e}_1\|} \right\}^2 = \left\{ \frac{\mathbf{a}^T \left( \sum_{j=1}^m b_{1j} \mathbf{b}_j \right)}{\|\mathbf{a}\| \left\| \sum_{j=1}^m b_{1j} \mathbf{b}_j \right\|} \right\}^2 \quad (8)$$

formát ölti.  $\|\mathbf{a}\| = 1$  természetesen, a képletben csak a világosabb megértés végett szerepeltettük. A teszt ugyanis, mint a formulából látható, valójában egy szög koszinuszának a négyzete, az  $\mathbf{a}$  vektor, valamint a  $\mathbf{s}_1 = \left( \sum_{j=1}^m b_{1j} \mathbf{b}_j \right)$  szabály által előállított vektor szögének koszinusznégyzete.

**58. Megjegyzés.** Az  $\|\mathbf{s}_1\| \geq \|\mathbf{s}_i\|$   $i = 1, \dots, n$  feltételi korlát alapján választjuk  $\mathbf{s}_1$ -et, ez felel meg az  $(\mathbf{x}_1 - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_1 - \bar{\mathbf{x}}) \geq (\mathbf{x}_i - \bar{\mathbf{x}})' \mathbf{A}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$  kritériumnak. Megfogalmazhatjuk úgy is, hogy  $\mathbf{B} \mathbf{B}^T$  legelső diagonális elemének kell a legnagyobb lenni.

<sup>7</sup> $m = 10$ ,  $n = 5000$  esetén például egy 50000 változós feladat adódik!

A (7) képlet szerint a  $[\mathbf{b}_j]$  rendszer ortonormált, az általa kifeszített altér pedig merőleges az összegző vektorra (ld. (6) képlet). A 5. feladatot tehát a következőképpen fogalmaztuk át: Vegyünk egy, az  $\mathbf{1}$  vektorra merőleges  $n-1$  dimenziós altérből egy ortonormált bázist, és a bázisvektorokból  $m$  darabot kiválasztva állítsuk elő azt az  $\mathbf{s}_1$ -et, mely  $\mathbf{a}$ -val a legkisebb szöget zárja be. Ismert matematikai tétel, hogy bármely ortonormált bázisból eljuthatunk bármely másik ortonormált rendszerbe ortogonális transzformációval, vagyis forgatással.

**59. Definíció.** Egyszerű forgatásnak nevezzük az olyan valódi ortogonális transzformációt, mely kétdimenziós síkbeli forgatást jelent, de változatlanul hagyja a síkra merőleges  $(n-2)$  dimenziós alteret, ennek mátrixa pedig az alábbi alakra hozható:

$$\mathbf{T}_i = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \cos \alpha_i & -\sin \alpha_i & \\ & & & \sin \alpha_i & \cos \alpha_i & \\ & & & & & 1 \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{bmatrix}$$

ahol az  $i$  index az  $\mathbf{e}_i, \mathbf{e}_{i+1}$  egységvektorok által kifeszített síkban történő forgatásra utal.

**60. Megjegyzés.** Azokat az ortogonális transzformációkat nevezzük valódinak, melyek determinánsa  $+1$ .

Egy  $(n-1)$  dimenzióban történő forgatást több egymást követő egyszerű forgatással tudunk végezni, egészen pontosan  $n-2$  darabbal. Mivel a feladatban szereplő  $n-1$  dimenziós altérnek az összegző vektorra kell ortogonálisnak lenni, ezért mindenekelőtt a triviális  $(\mathbf{e}_1, \dots, \mathbf{e}_{n-1}, \mathbf{e}_n)$  bázisról át kell térnünk egy olyan másik bázisra, melynek egyik eleme, például az  $n$ -dik éppen  $\frac{1}{\|\mathbf{1}\|}$  (vagyis  $\mathbf{e}_n$  a normalizált összegző vektorba megy át). így az első  $n-1$  bázisvektor által kifeszített altér erre lesz merőleges. Az előbbiekben mondottak szerint ez forgatással elérhető, legyen a forgatás mátrixa  $\mathbf{T} \equiv \mathbf{T}_1 \times \dots \times \mathbf{T}_{n-1}$ . Most már a megfelelő altérben vagyunk, a következő lépésben ennek bázisát kell oly módon elforgatni, hogy az átfogalmazott feladat optimális megoldását adja. Az előbbiek szerint ehhez már csak  $n-2$  egyszerű forgatásra van szükség, ennek mátrixa legyen  $\mathbf{Q} \equiv \mathbf{Q}_1 \times \dots \times \mathbf{Q}_{n-2}$ . A keletkezett  $\mathbf{QT}$  rendszerből kell  $m$  darabot kivenni, ehhez vezessük be az  $\mathbf{S} = [\mathbf{e}_{i_1}, \dots, \mathbf{e}_{i_m}]$  szelekciós mátrixot, ahol  $i_j \in \{1, \dots, n-1\}$ ,  $j = 1, \dots, m$ . A (8) képletbe helyettesítve tehát  $\mathbf{B} = \mathbf{QTS}$ .

A felírásban  $\mathbf{T}$  és  $\mathbf{S}$  konstans,  $\mathbf{Q}$  pedig  $n-2$  változó,  $\alpha = [\alpha_1, \dots, \alpha_{n-2}]$  függvénye. Az ortonormált rendszerek különleges struktúrájából adódóan

így ez a feladat csak  $n-2$  dimenziós (szemben a korábbi  $n \times m$ -es mérettel), és független attól, hogy  $m$  mekkora. Ez legrosszabb esetben is 4998 változót jelent, melyet a mai számítási kapacitások mellett már meg lehet oldani. Megjegyezzük, hogy a (6) és (7) képletek által megfogalmazott normalizálást a korábbi direkt megközelítésnél el kellett végezni (vagy előírni feltételi korlátként), itt azonban - értelemszerűen - nincsen rá szükség, így a számítási igény tovább csökken.

A  $\mathbf{T}$  forgatás meghatározása viszonylag egyszerű, általános alakja az alábbi:

$$\mathbf{T} = \begin{bmatrix} \cos \gamma_1 & -\cos \gamma_2 \sin \gamma_1 & \cdot & \pm \cos \gamma_{n-1} \prod_{i=1}^{n-2} \sin \gamma_i & \mp \prod_{i=1}^{n-1} \sin \gamma_i \\ \sin \gamma_1 & \cos \gamma_2 \cos \gamma_1 & \cdot & \mp \cos \gamma_{n-1} \cos \gamma_1 \prod_{i=2}^{n-2} \sin \gamma_i & \pm \cos \gamma_1 \prod_{i=2}^{n-1} \sin \gamma_i \\ 0 & \sin \gamma_2 & \cdot & \pm \cos \gamma_{n-1} \cos \gamma_2 \prod_{i=3}^{n-2} \sin \gamma_i & \mp \cos \gamma_2 \prod_{i=3}^{n-1} \sin \gamma_i \\ 0 & 0 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \sin \gamma_{n-1} & \cos \gamma_{n-1} \\ 0 & 0 & \cdot & 0 & 0 \end{bmatrix}$$

Ebből először  $\gamma_1$ -et tudjuk kiszámolni a  $\cos \gamma_1 + \sin \gamma_1 = 0$  egyenlet megoldásával, majd pedig  $\gamma_2$ -öt a  $\cos \gamma_2(\cos \gamma_1 - \sin \gamma_1) + \sin \gamma_2 = 0$  egyenletből és így tovább.

Mielőtt nekilátnánk a maximum kiszámolásának, érdemes megvizsgálni egy további lehetőséget.

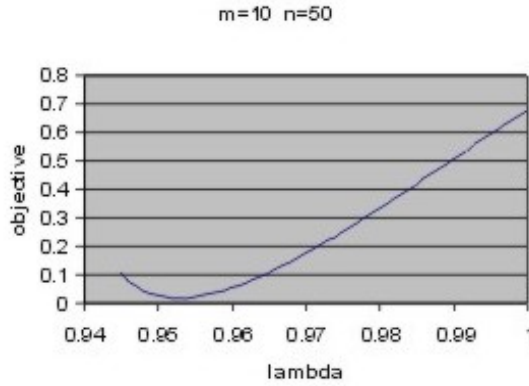
### 4.3. Egy másik megközelítés

Az előző pontokban azzal a hipotézissel éltünk, hogy a megfelelő normalizáló transzformáció  $\ln(W_{\max} - W)$  alakú, s ennek megfelelően megkíséreltük a formulában szereplő maximumot kiszámítani. Ne felejtsük el azonban, hogy annak mértéke, hogy egy közelítés mennyire jó, a Monte-Carlo szimuláció útján nyert empirikus eloszlásokon alapszik, azaz, hogy mennyire jól tudjuk a számunkra releváns percentiliseket közelíteni. Éppen ezért nem feltétlen érdemes ragaszkodni ahhoz a feltevéshez, hogy  $W_{\max}$  a Shapiro-Wilk teszt maximuma legyen, ha ezen hipotézistől függetlenül tudunk találni egy olyan  $\log(\lambda - W)$  alakú közelítést, amelynél az illeszkedés kellően pontos vagy esetleg pontosabb, és egyszerűen megvalósítható.

Tekintsük tehát a

$$\min_{\lambda} \sum [z_{\alpha_i}(\lambda) - \zeta_i]^2$$

$$\lambda \in (W_{10.000}, 1.0]$$



1. ábra. A célfüggvény értékének alakulása az  $m = 10$ ,  $n = 50$  esetben

optimalizálási feladatot, ahol  $\zeta_i$ -k a standard normális eloszlásból vett pontok alkalmasan megválasztott percentiliseknél (a célfüggvény paraméterei),  $z_{\alpha_i}(\lambda) = \left[ -|\log(\lambda - W_{\alpha_i})|^{1/2} - \mu \right] / \sigma$  pedig adott  $\lambda$  paraméter melletti közelített normális eloszlású változó, vagyis az  $\zeta_i$  pont visszabecslése.<sup>8</sup> Az optimum kritériuma tehát az, hogy az egyes percentiliseknél a becsült és tényleges normális eloszlású statisztika négyzetes eltérése minimális legyen.

Attól függően, hogy milyen normalizálást alkalmazunk, a szélsőérték feladatnak elvben létezhet analitikus megoldása. A numerikus vizsgálatok alapján nekünk a fenti normalizálás tűnt a legpontosabbnak, ez esetben azonban nem létezik analitikus megoldás. A numerikus megoldás azonban nem jelent komoly számítási nehézséget, mivel a feladat egyváltozós, a célfüggvény pedig - az empirikus vizsgálatok szerint - konvex.

Legyen  $\alpha_i \in \{0.01, 0.02, 0.05, 0.1, 0.5, 0.9, 0.95, 0.99\}$ .

A 1. ábrán a célfüggvény alakulását látjuk a  $\lambda \in (W_{10.000}, 1.0]$  tartományban egy találmány kivett esetben. Klasszikus optimalizálási módszerek használatával (Conn et.al. [1992]) előállítottuk tehát a  $\lambda^*$  értékek halmazát.

Ezek után három paramétert kellett regresszióval becsülnünk:  $\mu$ ,  $\sigma$  és  $\lambda$  értékét. Egy darab háromváltozós ( $m$  és  $n$  függvényében felírt  $\mu$ ,  $\sigma$  és  $\lambda$  változók) regresszióval nem sikerült kielégítő eredményt kapnunk. Tapasztalataink alapján a háromváltozós esetben nem lehet jó illeszkedést elérni, különösen az adatsorok szélein gyengék a közelítések. Ezért úgy döntöttünk, minden egyes  $m = \{2, \dots, 15\}$  dimenziószámra külön regressziót illesztünk, a mintaméreteket pedig az  $n_1 = \{(2m + 3) - 100(1)\}$ ,  $n_2 = \{110 - 200(10), 250 - 1000(50)\}$ ,  $n_3 = \{1500 - 5000(500)\}$  tartományokra

<sup>8</sup>A  $-\log(\lambda - W_{\alpha_i})^{1/2}$  transzformáció pontosabbnak bizonyult, mint a  $\log(\lambda - W_{\alpha_i})$  transzformáció, ezért végül is az előbbit használtuk.

bontottuk. Tehát 42 különböző regressziót kellett illesztenünk minden egyes paraméter becsléséhez, ami azonban praktikusán implementálható, mivel mindegyikük elsőfokú<sup>9</sup>. Magukat a kapott foks számokat és koefficienseket szükségtelennek tartottuk felsorolni, viszont elvégeztük a kapott regressziók felhasználásával a  $p$ -értékek visszabecslését. A becslések abszolút hibáit, vagyis a  $\Phi(z_{\alpha_i}(\lambda)) - \alpha_i$  különbségeket elemeztük, a pontosság kritériuma az volt, hogy 1, 5 illetve 10% esetén a hiba ne érje el a fél százalékpontot, míg 1 – 5 ezrelék esetén a fél ezreléket (bár utóbbi nem túl érdekes gyakorlati problémáknál). Az előjelek alapján arra következtethettünk, hogy nincs szisztematikus hiba a becslésekben. Megjegyezzük, hogy a percentilisek, melyeken a becslés alapszik, szintén tartalmazznak némi hibát, bár kicsit, hiszen  $r = 10.000$  méretű Monte-Carlo szimulációból származnak.

#### 4.4. Az egyes megközelítések értékelése

Az eredmények szerint  $\lambda^* \sim W_{\max}$ , legalábbis egy bizonyos  $m$ -től kezdve. Kisebb  $m$ -ekre a  $\lambda$  optimális értéke igen közel esik 1-hez a legtöbb mintaméret esetén (míg  $W_{\max}$  kis mintáknál kisebb 1-nél). Úgy tűnik tehát, hogy a  $\min \sum [z_{\alpha_i}(\lambda) - \zeta_i]^2$  feladat megoldásával kapott közelítés stabilabb és pontosabb közelítés, emellett kiszámítása nem kíván különösebb gépi erőforrást.

1. táblázat. Abszolút hibák,  $m=10$

$n =$	0.001	0.005	0.01	0.05	0.1	$W_{10.000} + \epsilon$
25	0.001	0.005	0.006	0.007	0.008	0.87080765
30	0.001	0.003	0.004	0.007	0.006	0.90064373
35	0	0.001	0.003	0.007	0.005	0.91930916
40	0.001	0.002	0.004	0.011	0.01	0.92727591
50	0	0.003	0.005	0.006	0.01	0.94490731

2. táblázat. Abszolút hibák,  $m=10$

$n =$	0.001	0.005	0.01	0.05	0.1	$\lambda^*$
25	0	0.001	0	-0.003	-0.002	0.90084653
30	0	0	-0.001	-0.001	-0.002	0.91997325
35	0	0	0	0.001	-0.001	0.92779875
40	0	0	0	0.003	0.002	0.93881353
50	0	0	0.001	-0.001	0.002	0.95288836

A 1. és 2. táblázatokban közöljük egy tetszőlegesen kiválasztott esetben a kétféle megközelítés alapján számított abszolút hibákat.

<sup>9</sup>Ez persze nem jelenti azt, hogy lineárisak is, a törtkitevős polinomokra adott foks szám definíció alapján elsőfokúak.

Ha  $W_{\max}$  egzakt értékét meg tudnánk határozni, talán érdemes lenne a megközelítést továbbvinni, hiszen egy analitikus megoldást a pontosság érdekében preferálnánk egy numerikusan meghatározott, majd pedig polinommal közelített megoldással szemben. Mivel azonban  $W_{\max}$  pontos értékét nem tudjuk megadni (legalábbis egyelőre), így nem indokolt egy  $n-2$  méretű numerikus feladattal foglalkozni, ha létezik pontosabb, egyváltozós alternatíva.

## 5. A Shapiro-Wilk teszt egyéb többdimenziós kiterjesztései

A Shapiro-Wilk-féle teszt többdimenziós kiterjesztésére számos variáns készült. A következő alpontot ezek ismertetésének szenteljük. Ezután vizsgálatot végzünk a próbák erejére vonatkozóan, bevonva a vizsgálódásba az általunk tárgyalt *Malkovich-Affifi* által javasolt változatot.

### 5.1. Rövid áttekintés

A Shapiro-Wilk teszt többdimenziós kiterjesztései szinte kivétel nélkül azt az elvet követik, hogy valamilyen alkalmas transzformációval vagy dekompozícióval az alapeloszlást visszavezetjük egy vagy több egyváltozós problémára.

A *Malkovich-Affifi* [1973] által javasolt kiterjesztést már ismertettük az előző pontokban részletesen. Itt az  $y = \mathbf{c} \cdot \mathbf{x}$  lineáris transzformációval képeztünk egydimenziós eloszlást, és erre alkalmaztuk a Shapiro-Wilk tesztet.

A *Mudholkar et.al.* [1995] tanulmányban a szerzők az  $m$ -dimenziós eloszlást felbonják  $m$  darab egydimenziós eloszlásra, majd kiszámítják ezek szignifikancia szintjét, vagyis a  $p$ -értékeket. A felbontás a következőképpen történik. Legyen

$$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n] \quad \mathbf{x}_i \in \mathbb{R}^m \quad (9)$$

egy  $m$ -dimenziós normális eloszlásból vett  $n$  elemű minta, valamint vezessük be az  $\mathbf{B} = \mathbf{X}^T = [\mathbf{b}_1, \dots, \mathbf{b}_m]$   $\mathbf{b}_j \in \mathbb{R}^n$  jelölést. Legyen továbbá  $\mathbf{b}^{(k)} = [\mathbf{1}, \mathbf{b}_1, \dots, \mathbf{b}_{k-1}]$  és becsüljük a

$$\mathbf{b}_k = \beta^{(k)} \mathbf{b}^{(k)} + \epsilon^{(k)} \quad k = 2, \dots, m$$

regressziókat. Ekkor  $\epsilon^{(k)} \sim N_n \left( \mathbf{0}, \sigma_{k|12\dots(k-1)}^2 \mathbf{I}_n \right)$ , ahol  $\sigma_{k|12\dots(k-1)}^2$  a feltételes variancia. Létezik olyan  $(n-k) \times n$  méretű  $\mathbf{G}_k$  mátrix, hogy  $\mathbf{G}_k \mathbf{G}_k^T = \mathbf{I}_{n-k}$  és  $\mathbf{G}_k \mathbf{b}^{(k)} = \mathbf{0}$ . Most legyen

$$\mathbf{b}_k^* = \mathbf{G}_k \mathbf{b}_k \quad (10)$$



Ekkor  $\mathbf{b}_k^*$ -nak  $\mathbf{b}^{(k)}$ -ra vonatkozó feltételes eloszlása  $N_{n-k} \left( \mathbf{0}, \sigma_k^2 |_{12\dots(k-1)} \mathbf{I}_{n-k} \right)$ , vagyis  $\mathbf{b}_k^*$  független  $\mathbf{b}_1, \dots, \mathbf{b}_{k-1}$ -től és így független  $\mathbf{b}_2^*, \dots, \mathbf{b}_{k-1}^*$ -től is. Megfogalmazható a következő állítás:

**61. Állítás.** A (9) minta többdimenziós normalitásának tesztelése ekvivalens az  $m$  darab, a (10) által meghatározott  $\mathbf{b}_1^* = \mathbf{b}_1, \mathbf{b}_2^*, \dots, \mathbf{b}_m^*$  egydimenziós minták normalitásának tesztelésével.

A teszt elvégzéséhez illetve  $p$ -értékek kiszámítására természetesen a Royston-féle (Royston [1995]) algoritmust használtuk. Igaz a következő állítás:

**62. Állítás.** A  $\mathbf{b}_1^*, \mathbf{b}_2^*, \dots, \mathbf{b}_m^*$  mintákból számított  $W$  statisztikák, valamint az ehhez tartozó  $p_1, \dots, p_m$  értékek függetlenek.

A kapott  $p$ -értékek többféleképpen is kombinálhatók. Az egyes megközelítések a következők:

**F** Fisher-féle teszt:  $W_F = -2 \sum \log p_j$

**L** Logit statisztika:  $W_L = A^{-1/2} \sum \log(p_j / (1 - p_j))$ ,  $A = \pi^2 m (5m + 2) / (15m + 12)$

**N** Lipták-féle teszt:  $W_N = \sum \Phi^{-1}(1 - p_j)$

**T** Tippett-féle teszt:  $W_T = \min p_j$

Maga Royston is javasolt többdimenziós Shapiro-Wilk tesztet (Royston [1983]). Lényege abban áll, hogy komponensenként az egydimenziós  $W$ -teszt formuláját alkalmazva olyan változókhoz juthatunk, melyek közt kicsi a korreláció, függetlenül attól, hogy az alapeloszlásban az egyes változók mennyire voltak korreláltak. A 3.1. pontban ismertettük, hogy a  $W$ -statisztikából közelítőleg normális eloszlású változót tudunk transzformálni, ahol a  $\mu$  várható értéket illetve  $\sigma$  szórását regressziós becsléssel állítjuk elő a mintaméretből ( $n$ ).

Számítsuk ki a  $z_1, \dots, z_m$  statisztikákat az előbb leírt módon, ahol  $z_j$  az eloszlás  $j$ -dik változójából származik, majd képezzük a

$$k_j = \left\{ \Phi^{-1} \left[ \frac{1}{2} \Phi(-z_j) \right] \right\}^2 \quad j = 1, \dots, m$$

statisztikákat. Ha a mintabeli változók korrelálatlanok volnának, akkor a belőlük számított  $W$ ,  $z$  és  $k$  statisztikák is azok lennének. Ekkor

$$G \equiv \frac{1}{m} \sum_{j=1}^m k_j$$

$\chi_m^2/m$  eloszlást követne. Ha pedig a korreláció egységnyi lenne, akkor  $k_1 = k_2 = \dots = k_m$  miatt  $G \sim \chi_1^2$ .

Gyakorlati szempontból nyilvánvalóan a köztes eset érdekes, amit egy  $G \sim \chi_e^2/e$  statisztikával jellemezhetnénk, ahol  $1 < e < m$  nem feltétlen egész szám.

**63. Megjegyzés.** Más szavakkal a peremekből nyert  $z_j$  statisztikák  $m$ -dimenziós normális eloszlást követnek, ha függetlenek, és egydimenziós (degenerált) eloszlást, ha egységnyi a korreláció. A köztes eseteket úgy lehetne felfogni, mintha a normális eloszlást tört-dimenziókra kiterjesztenénk.

Az  $e$  értéke a következőképpen származtatható az első- és másodrendű momentumokból:

$$\begin{aligned}\mu_G &= 1 \\ m^2 \sigma_G^2 &= 2m + \sum_{j \neq i}^m \sum_{i \neq j}^m \text{cov}(k_i, k_j)\end{aligned}$$

Mivel  $\sigma_{k_j}^2 = 2$  és  $c_{ij} \equiv \text{corr}(k_i, k_j) = \text{cov}(k_i, k_j) / \sqrt{2 \times 2}$  ezért

$$\begin{aligned}m^2 \sigma_G^2 &= 2m + 2 \sum_{j \neq i}^m \sum_{i \neq j}^m c_{ij} \\ \sigma_G^2 &= \frac{2}{m} + \frac{2}{m^2} \sum_{j \neq i}^m \sum_{i \neq j}^m c_{ij}\end{aligned}$$

Ugyanakkor  $\sigma^2(\chi_e^2/e) = 2e/e^2 = 2/e$ , így

$$\begin{aligned}2/e &= \sigma_G^2 \\ e &= \frac{2}{\sigma_G^2} = \frac{1}{\frac{1}{m} + \frac{1}{m^2} \sum_{j \neq i}^m \sum_{i \neq j}^m c_{ij}}\end{aligned}$$

Vagyis a kombinált többdimenziós normalitásteszt ( $H$ -teszt a továbbiakban) a

$$H \equiv eG = \frac{e}{m} \sum_{j=1}^m k_j \sim \chi_e^2$$

formájú.

**64. Megjegyzés.**  $e = m$  ha  $c_{ij} = 0$   $i, j = 1, \dots, m$  és  $i \neq j$ , valamint  $e = 1$ , ha  $c_{ij} = 1$ .

A  $c_{ij}$  értékek pedig a következő módon becsülhetők:

$$\hat{c}_{ij} = \begin{cases} r_{ij}^\lambda \left[1 - \frac{\mu}{\nu} (1 - r_{ij})^\mu\right], & i \neq j \\ 1, & i = j \end{cases}$$

$r_{ij}$  a mintából számított páronkénti korrelációs együttható, a többi paraméterre pedig a  $\lambda = 5$ ,  $\mu = 0.715$  és  $\nu = 0.35$  közelítések használhatók  $n = 2000$  mintaméretig (Royston [1983]).

Ha az ismertett Shapiro-Wilk-féle többdimenziós variánsokat osztályozni szeretnénk, akkor a  $H$ -teszt igen közel áll 1.2. pontban ismertett típusok közül a  $T1$ -hez. A peremekre vonatkozó információk alapján próbál ugyanis az együttes normalitásra következtetni, és a teszt statisztika megkonstruálásához az egyes változók közti lineáris korrelációt igyekszik kiszűrni. Ezzel szemben az  $F$ -,  $L$ -,  $N$ - és  $T$ -tesztek a lineáris regresszió illetve a lineáris korreláció fogalmán keresztül próbálják az együttes normalitás jellemzőit megragadni, így besorolhatjuk őket a  $T2$  csoportba. A  $W$ -teszt definíció szerint a  $T3$  csoporthoz tartozik, hiszen a többdimenziós eloszlás egy kitüntetett metszetének vizsgálatára épül. Történelmileg szerencsésen alakultak tehát a tények, ugyanarra az egydimenziós tesztre vonatkozó három olyan kiterjesztést tudunk megvizsgálni, melyek három különböző megközelítés valamelyikéhez tartoznak.

## 5.2. A próbák erejének vizsgálata

A próbák erejének vizsgálatához a Mudholkar et.al. [1995] cikket vettük alapul. A szerzők különféle két és három dimenziós eloszlásokból vettek 20 – 50 elemű mintákat, és vizsgálták az elutasítás mértékét különböző tesztek esetében  $\alpha = 10\%$ ,  $5\%$  és  $1\%$  szignifikancia szinteken. Konkrétan a következő eloszlásokról van szó:

**Többdimenziós  $\chi^2$  eloszlások:** Legyenek  $x_1, x_2, \dots, x_m$  független  $\chi^2$  eloszlású változók  $\nu_1, \nu_2, \dots, \nu_m$  szabadságfokokkal, valamint legyen  $x$  egy további független  $\chi^2$  eloszlású változó  $\nu$  szabadságfokkal. Ekkor  $y_j = x + x_j$  változók függetlenek, és a peremeloszlásuk  $\nu + \nu_j$  szabadságfokú  $\chi^2$  eloszlás. Jelölése:  $\chi_m^2(\nu_1, \dots, \nu_m; \nu)$

**Többdimenziós Cauchy és  $t$ -eloszlás:** Az  $m$ -dimenziós  $\nu$  szabadságfokú  $t$ -eloszlású véletlen számot az alábbi transzformációval képezzük:

$$\mathbf{y} = \frac{\mathbf{z}}{s_\nu / \sqrt{\nu}} + \mu$$

ahol  $s$   $\nu$  szabadságfokú  $\chi$  eloszlású véletlen szám,  $\mathbf{z}$  ettől független,  $N_m(\mathbf{0}, \mathbf{I})$  eloszlásból nyert vektor,  $\mu$  pedig konstans, az eloszlás várható értéke.  $\nu = 1$  esetén értelemszerűen többdimenziós *Cauchy*-eloszlást kapunk.

Kever normális eloszlás: Kever normális eloszlású véletlen számokat a

$$\mathbf{y} = \mu + \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2$$

transzformációval kapunk, ahol  $\mathbf{x}_1 \sim N_m(\mathbf{0}, \Sigma)$ ,  $\mathbf{x}_2 \sim N_m(\mathbf{0}, \gamma^2 \Sigma)$ ,  $\gamma \geq 0$ , és  $\lambda \sim Be$ , vagyis  $\Pr(\lambda = 1)$  valószínűséggel az egyik,  $\Pr(\lambda = 0)$  valószínűséggel pedig a másik eloszlásból vettünk mintát.

Többdimenziós lognormális eloszlás: ebből az eloszlásból úgy generálunk véletlen számokat, hogy először mintát veszünk egy  $N_m(\mathbf{0}, \Sigma)$  eloszlásból, majd koordinátáinként képezzük az  $y_j = e^{x_j}$ ,  $j = 1, 2, \dots, m$  transzformációt.

Többdimenziós Burr-Pareto-Logisztikus eloszlás: Vegyünk először egy független  $m$  elemű,  $x_1, x_2, \dots, x_m$  mintát exponenciális eloszlásból, majd pedig ettől függetlenül egy  $v$  elemet egy 1 skála paraméterű gamma eloszlásból. Ekkor

$$y_j = (1 + x_j/v)^{-\theta}$$

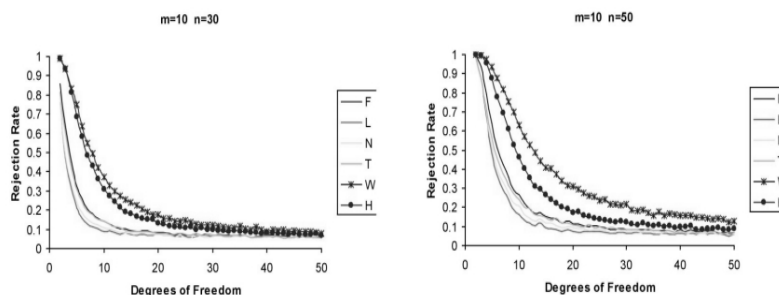
az  $m$ -dimenziós Burr-Pareto-Logisztikus eloszlású véletlen változó vektor egyes koordinátáit jelöli,  $\theta$  alkalmas forma tényező.

$\Omega_m(\mu, \Sigma, \nu)$ -eloszlás: ez egy általunk definiált eloszlás, mely  $\mu_j$  és  $\sigma_j$  paraméterű normális eloszlású peremekkel rendelkezik, a függőségi struktúráját pedig egy  $(\mathbf{R}, \nu)$  paraméterű  $t$ -kopula határozza meg, ahol  $\sigma_{ij} = r_{ij}\sigma_i\sigma_j$ . A minta vétel úgy történik, hogy generálunk a megadott paraméterű  $t$ -kopulából  $x_1, x_2, \dots, x_m$  véletlen számokat, majd képezzük az  $y_j = \Phi^{-1}(x_j)$  transzformációkat, ahol  $\Phi^{-1}(\cdot)$  a normális eloszlás inverz eloszlásfüggvénye.

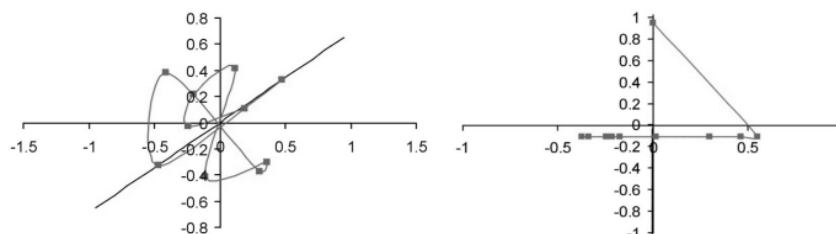
**65. Megjegyzés.** Intuitíve könnyen belátható, hogy  $\Omega_m(\cdot)$  elliptikus eloszlás.

A cikkben az  $m = 2, 3$  eseteket vizsgálták. Ezeket az eseteket reprodukáltuk, de a szimulációt kiterjesztettük a minket érdeklő  $m = 2, \dots, 15$  tartományra. A vizsgálatba ugyanakkor bevontuk e tanulmány tárgyát képező *Malkovich-Affifi-féle* variánst is, amit az egyszerűség kedvéért a továbbiakban  $W$  tesztként említünk, valamint a korábban ismertetett Royston-féle  $H$ -tesztet is. A mintaméret továbbra is az  $n = 20 - 50$ , hiszen a próbák erejének vizsgálata mindig a kis mintákra érdekes.

Terjedelmi korlátok miatt néhány számszerű eredményt mutatunk be az *I. Függelék*-ben, a fontosabb következtetéseinket grafikonokkal illusztráljuk. A vizsgálat igen meglepő eredménnyel szolgált. Míg az  $F$ ,  $L$ ,  $N$  és  $T$ -tesztek hasonló eredményeket produkáltak, addig a  $W$  variáns általában jelentősen elmaradt mögöttük, kivéve két esetet: a Student  $t$ -eloszlást és a kevert



2. ábra. Próbák erejének alakulása  $t$ -eloszlás esetén



3. ábra.

Kétváltozós normális eloszlásból vett minták, melyek a  $W = W_{\max}$  és  $W = W_{\min}$  értékeket adják

normális eloszlást. Ami a  $t$ -eloszlást illeti, a helyzet éppen a fordítottja. A  $H$ -teszt viszont általában jelentősen erősebbnek bizonyult, mint bármely másik. A  $t$ -eloszlás esetében ez a teszt is jelentősen felülmúlta a többi, a  $W$ -teszt azonban, úgy tűnik, itt abszolút mértékben a legjobb.

Illusztráció céljából találomra kivettünk két esetet, és megnéztük, hogyan alakul a próbák ereje mondjuk  $\alpha = 5\%$  szignifikancia szintnél, ahogy a  $t$ -eloszlás szabadságfokát növeljük, vagyis tartunk a normális eloszláshoz.

Tekintsük a 2. ábrát! A vízszintes tengely a szabadságfokokat jelöli, a  $\nu = 2, \dots, 50$  tartományt ábrázolja egyes lépésközökkel, míg a függőleges tengely az elutasítás valószínűségét jelöli. A  $*$ -mintázatú görbe jelöli a Malkovich-Affifi-féle tesztet, a  $\bullet$ -mintázatú a  $H$ -tesztet, míg a sima vonalak a másik négyet. Az ábrák önmagukért beszélnek, a  $W$ -teszt ereje jelentősen nagyobb a  $t$ -eloszlás esetében, mint a másik öté. Ez azért érdekes, mert a  $t$ -eloszlás igen hasonló a normális eloszláshoz, és magasabb dimenziókban éppen az okozza a problémát, hogy általában gyengék a tesztek.

Hogy miért jobb a  $W$ -teszt a többinél, az intuitíve nagyon könnyen belátható.

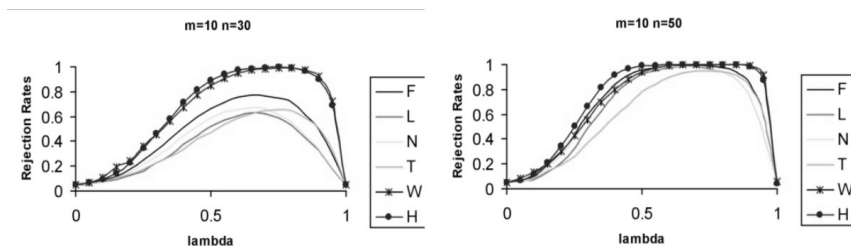
Az 3. ábrán két  $m = 2$  dimenziós, standard normális eloszlásból származó mintát ábrázoltunk. A bal oldali grafikon azt a mintát ábrázolja, ahol a  $W$ -tesztnek maximuma van, a jobb oldali pedig egy olyan mintát, ahol minimuma van. A bal oldali egy tipikus normális eloszlású minta, nagyjából körkörösén helyezkednek el a mintaelemek, az origóból kifelé haladva viszonylag egyenletesen csökken a mintába bekerülő elemek valószínűsége (bár ilyen kis mintánál ez az ábrán nemigen látható). A jobb oldali egy extrémabb, de nyilván valós eset, ahol egy mintaelem kiugró, míg a többi ehhez képest sűrűbben helyezkedik el.

Az ábrán az egyes mintaelemeket vonallal kötöttük össze, mégpedig a normájuk nagysága szerint. Standard normális eloszlás esetén a  $W$  statisztika képzése geometriai értelemben azt jelenti, hogy az egyes mintaelemeket a legnagyobb normájú mintaelem és az origó által meghatározott egyenesre, mint alátérre képezzük le (a többdimenziós felület egy egydimenziós szeletét vizsgáljuk, ld. 1.2. rész). A megfelelő altereket szintén szerepeltettük az ábrákon.

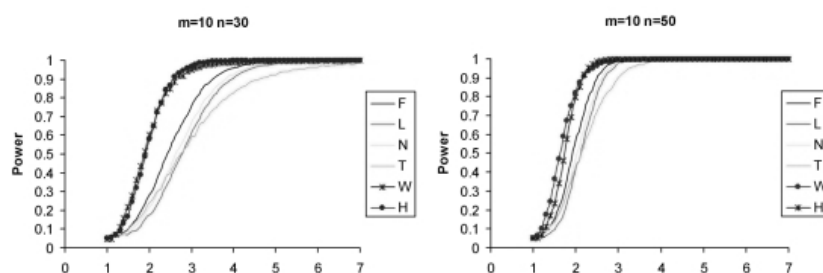
Az előbbieket szerint tehát ha van a mintában kiugró elem, akkor mindig az általa meghatározott alátérbe történik a leképzés. Az ilyen típusú minták pedig a vastag farkú eloszlások, így a  $t$ -eloszlás esetében sokkal jellemzőbbek, mint a normális eloszlásnál. Több kiugró mintaelem esetén pedig — különösen magasabb dimenziókban — igen csekély annak az esélye, hogy ezek mind egy egyenesre esnek, tehát az egydimenziós esetben megszokott szimmetria elvész. A leghosszabb normájú elem a többi kiugró elem elhelyezkedését a leképzés során vélhetően jelentősen „torzítani” fogja. Ez egy igen érdekes, és feltehetően egy értékes tulajdonsága ennek a változatnak.

A kevert eloszlások esetében a  $W$  statisztika egyértelműen dominálja a többi variánst. A keverés alapjául szolgáló egyik eloszlás  $\Sigma$  variancia-kovariancia mátrixszal rendelkezik, a másik pedig egy  $\gamma\Sigma$  mátrixszal, ahol  $\gamma \geq 1$ . A korrelációs mátrixuk tehát azonos, és ez lesz a kevert normális eloszlás korrelációs mátrixa is, a keverés alapjául szolgáló két normális eloszlás varianciája különbözik. A 31. állítás alapján tudjuk, hogy ez az eloszlás elliptikus, így a korrelációs mátrix jelentése azonos a normális eloszlásnál használt fogalommal. Ha  $\gamma = 1$ , akkor a normális eloszlást kapjuk vissza,  $\gamma > 1$  értékekre pedig egyre távolodunk a normális eloszlástól. A normalitástól való eltérés mértéke ugyanakkor a keverési aránytól is függ. Ha az egyik eloszlásból vesszük a minta jó részét, mondjuk 90%-át, akkor a normális eloszláshoz megintcsak hasonlatos lesz a kapott eloszlás.

Az empirikus eredmények alapján a korreláció egyáltalán nem befolyásolja az  $F$ ,  $L$ ,  $N$ ,  $T$  és  $W$ -tesztek erejét, ezért a  $\Sigma = \mathbf{I}$  és  $\gamma = 2$  beállítás mellett végeztük el a vizsgálatot. A 4. ábrán ezt esetet ábrázoltuk,  $m = 10$  és  $n = 30, 50$  most is, a próbák erejét a keverési arány függvényében tüntettük fel. A kisebb,  $n < 30$  méretű mintáknál lehet nagyobb különbséget megfigyelni a tesztek között, a  $W$ -teszt dominanciája itt a legszembetűnőbb. Az is leolvasható az ábráról, hogy 50% körüli keverési aránynál a legnagyobb a



4. ábra. Próbák erejének alakulása kevert normális eloszlás esetén



5. ábra.

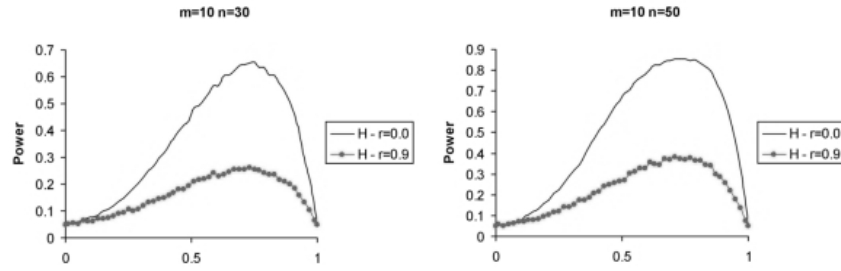
Próbák erejének alakulása a kiinduló két normális eloszlás távolságának függvényében

normalitástól való eltérés, melyet intuitíve is gondoltunk.

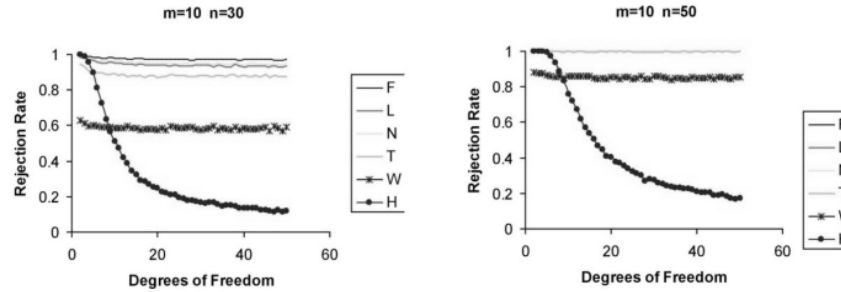
Ha  $\gamma$  értékét növeljük, akkor ez növeli a normalitástól való eltérést, és növekszik a próbák ereje. A  $W$ -teszt dominanciája relatíve csökken, hiszen a normalitástól való eltérést könnyebb felismerni. A 5. ábrán most a  $\gamma$  függvényében ábrázoltuk a próbák erejének alakulását egy fix, 60%-os keverési arány mellett. A  $\gamma = 1$  értéknél minden teszt esetében 5% körüli értéket becslünk, hiszen ez normális eloszlás. Ettől távolódva a  $W$ -teszt ereje a többiekéhez képest rohamosan növekszik, majd pedig magas  $\gamma$  értékekre a többiek is felzárkóznak. Levonhatjuk tehát azt a következtetést, hogy a  $W$ -teszt érzékenyebb ennél az eloszláscsaládnál, mint a többiek.

A korreláció nagyságára a tesztek invariánsak voltak, kivéve a  $H$ -tesztet. A 6. ábrán felrajzoltuk a  $H$ -teszten elvégzett vizsgálat eredményét  $r_{kj} = 0.0$  és  $r_{kj} = 0.9$  mértékű korrelációra. A korreláció növekedésével a teszt drasztikusan gyengül. Ennek az lehet egy intuitív magyarázata, hogy a közel degenerált eloszlások keverésekor a peremek sokkal kisebb mértékben módosulnak, mint különböző varianciájú, de független peremekkel rendelkező normális eloszlások keverésekor.

Érdeemes még nyomon követni a próbák erejének alakulását az  $m$ -dimenziós  $\chi^2$  eloszlás esetén, ha a peremek  $\chi_k^2$   $k = 3, \dots, 50$  szerint alakul-



6. ábra. A  $H$ -teszt erejének alakulása zérus és erős korreláció esetén



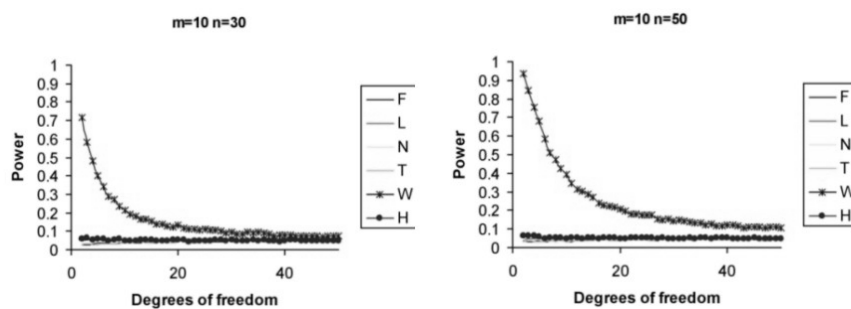
7. ábra. Próbák erejének alakulása  $m$ -dimenziós  $\chi^2$  eloszlás esetén

nak, itt látszik ugyanis a  $H$ -teszt gyengesége (ld. 7. ábra). A peremek a szabadságfok növekedésével kezdenek egyre inkább egy haranggörbéhez hasonlítani, a  $\chi^2$  eloszlás peremei közel szimmetrikussá válnak. Sem a  $W$ -tesztet, sem pedig az  $F - T$  teszteket ez nem különösebben befolyásolja, hiszen valódi többdimenziós karakterisztikákra építenek (az  $F$ ,  $L$ ,  $N$  és  $T$ -tesztek például a normalitás által implicált linearitásra), a  $H$ -teszt ugyanakkor a peremekre vonatkozó információkat összegzi egy tesztben. Ezért a  $H$ -teszt gyakorlatilag a 1.2. pontban felsorolt csoportból a  $T1$ -hez sorolható (nem valódi  $m$ -dimenziós teszt).

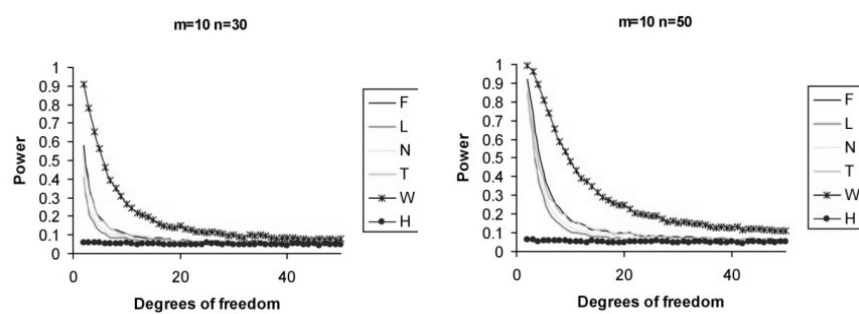
Végezetül nézzük meg az  $\Omega_m(\mu, \Sigma, \nu)$  eloszlásra vonatkozó eredményeket! A 8. ábra  $\mu = \mathbf{0}$ ,  $\Sigma = \mathbf{I}$  és  $\nu = 3, \dots, 50$  eseteket mutatja. Az eredmény igencsak meglepő! Míg a  $W$ -teszt nagyon hasonlóan viselkedik, mint egy  $t$ -eloszlás esetén, addig a többi tesztnél közelítőleg 5%-os elutasítási rátákat számoltunk, ami az elsőfajú hiba valószínűsége. Ez azt jelenti, hogy az eloszlást normális eloszlásnak érzékelik.

A 9. ábrán olyan esetet ábrázoltunk, amikor is  $\Sigma$ -ban  $\sigma_{jj} = 1$  és  $\sigma_{jk} = 0.9$ , vagyis erős lineáris korreláció van jelen. Az  $F$ ,  $L$ ,  $N$  és  $T$ -tesztek itt már érzékelik a normalitás hiányát, a  $H$ -teszt azonban továbbra sem. Mint tudjuk, a  $t$ -eloszlás peremváltozói sohasem függetlenek, még zérus korreláció esetén sem. Ezt a függőséget a  $H$ -teszt azonban egyáltalán





8. ábra. Normális eloszlású peremek és Student  $t$ –kopula, zérus korreláció



9. ábra. Normális eloszlású peremek és Student  $t$ –kopula, erős korreláció

nem képes érzékelni, vagyis a peremeket függetlennek ítéli. Ennek az lesz a következménye, hogy elfogadja az együttes normalitást, hiszen független normális eloszlású változók együttes eloszlása is normális. Az  $F$ ,  $L$ ,  $N$  és  $T$ -tesztek ezt a függést bizonyos szinten érzékelik, de csakis korreláltság esetén. Emlékezzünk ugyanis vissza, hogy ezeknek a teszteknek a konstrukciójában lineáris regressziókat használtunk, melyek viszont zérus korrelációnál nem működnek, ekkor ugyanis nincs lineáris kapcsolat.

**66. Megjegyzés.** *A próbák ereje tehát attól is függ, hogy az egyes variánsok mennyire érzékelik a peremek közti függőséget. Az  $F$ ,  $L$ ,  $N$  és  $T$ -tesztek csak a lineáris korreláción keresztül, annak mértékében képesek az adott eloszlásban lévő nemlineáris függőségi struktúrát felismerni, mely elliptikus  $H_1$ -esetén igen komoly hátrány. A  $H$ -teszt viszont elliptikus eloszlásokra csak akkor használható, ha a peremek jól megkülönböztethetők a normális eloszlástól.*

### 5.3. A valódi többdimenziós Shapiro-Wilk teszt

A 2.2.pont 53. megjegyzésében említettük, hogy a teszt kiszámítására adott eljárás valójában nem a  $\min_{\mathbf{c}} W(\mathbf{c})$  értéket, hanem az egydimenziós esetben a minimum helyhez négyzetes hibában legközelebbi ponton számítja ki a teszt értékét. A kettő viszont csak speciális feltételek (pl. monotonitás) mellett esik egybe.

Az előző pontban kapott eredményeink alapján úgy gondoltuk, érdekes lenne megnézni, vajon mennyiben változnak meg a teszt tulajdonságai, ha a tesztet a tényleges minimumban számítjuk ki. Azt reméljük ugyanis, hogy a teszt erejét javítani tudjuk majd. A teszt alkotóinak eredeti szándéka az volt, hogy a többdimenziós eloszlás metszetei közül a legrosszabb esetet tekintsük, ami annak felel meg, mikor a teszt értéke minimális. Mivel a megvalósított változat ténylegesen nem felel meg ennek az elgondolásnak, ezért intuitíve azt gondoljuk, a teszt viselkedése valószínűleg elég labilis. Adott minta szabja meg ugyanis a tesztnek, mint függvénynek az alakját, így mintáról mintára a számított érték illetve a tényleges minimum hol közelebb, hol távolabb van egymástól.

A kiszámításhoz jó kiinduló alapot szolgáltat az egydimenziós esetben követett eljárás (ld. 46. tétel). A különbség annyi, hogy nem  $y_i$ -ben, hanem  $\mathbf{c}$ -ben keressük a megoldást.

Legyen tehát  $\mathbf{x}_i \in \mathbb{R}^m$ ,  $i = 1, \dots, n$  ( $m+1 < n$ ) az  $m$ -dimenziós normális eloszlásból származó minta, és tegyük fel, hogy  $\sum \mathbf{x}_i = \mathbf{0}$ , vagyis normáljuk a mintát eszerint. Nézzük az alapformulát az  $y_i = \mathbf{c} \cdot \mathbf{x}_i$  helyettesítéssel:

$$W = \frac{\left( \sum_{i=1}^n a_i \mathbf{c} \cdot \mathbf{x}_i \right)^2}{\sum_{i=1}^n (\mathbf{c} \cdot \mathbf{x}_i)^2} = \frac{\left[ \mathbf{c} \cdot \left( \sum_{i=1}^n a_i \mathbf{x}_i \right) \right]^2}{\sum_{i=1}^n (\mathbf{c} \cdot \mathbf{x}_i)^2}$$

Mivel a formula  $\mathbf{c}$ -ben is nullad fokon homogén, ezért a

$$\mathbf{c}' \left( \sum_{i=1}^n a_i \mathbf{x}_i \right) = 1$$

szinten most is megköthetjük a számláló értékét. A teszt értéke akkor minimális, ha a nevező maximális. Mivel a tesztet rendezett statisztikákra kell kiszámítani, ezért az optimumban a  $\mathbf{c}'\mathbf{x}_i \leq \mathbf{c}'\mathbf{x}_{i+1}$  összefüggéseknek teljesülniük kell.

**67. Megjegyzés.** Amikor a 4.2.pontban a teszt maximumát szerettük volna  $\mathbf{x}$ -ben kiszámítani, akkor nem kellett a rendezés miatt aggódnunk, a maximális korrelációt ugyanis a rendezett minta adta. Mivel itt minimumra törekszünk, a rendezettség nyilvánvalóan egy effektív korlátot fog jelenteni.

Az optimum feladat formálisan így

$$\begin{aligned} \max_{\mathbf{c}} \sum_{i=1}^n (\mathbf{c}'\mathbf{x}_i)^2 & \quad (11) \\ \mathbf{c}'(\mathbf{x}_i - \mathbf{x}_{i+1}) & \leq 0 \quad i = 1, \dots, n-1 \\ \mathbf{c}' \left( \sum_{i=1}^n a_i \mathbf{x}_i \right) & = 1 \end{aligned}$$

Megjegyezzük, hogy a célfüggvény most is konvex, ugyanis

$$\begin{aligned} \frac{\partial^2 \sum_{i=1}^n (\mathbf{c}'\mathbf{x}_i)^2}{\partial c_j \partial c_k} & = 2 \sum_{i=1}^n x_{ij} x_{ik} \\ \frac{\partial^2 \sum_{i=1}^n (\mathbf{c}'\mathbf{x}_i)^2}{\partial c_j^2} & = 2 \sum_{i=1}^n x_{ij}^2 \end{aligned}$$

vagyis a függvény Hesse-mátrixa a mintából számított variancia-kovariancia mátrix (egészen pontosan annak  $2n$ -szerese), ami viszont  $m+1 < n$  esetén pozitív definit (1 valószínűséggel). A feltételek által megadott halmaz pedig egy konvex poliéder, így a maximum ennek valamelyik csúcsában éretik el.

Egy további problémát okoz azonban, hogy hogyan rendezzük az eredeti,  $m$ -dimenziós mintát! Egy minta illetve annak valamely permutáltja nyilván ugyanaz a minta. Mi azt a rendezést tekintettük mérvadónak, amelynél

1. a (11) feladatnak létezik megoldása, valamint
2. a legkisebb optimumot szolgáltatja

Az első közelítésben így  $n!$  esetet kellene végig néznünk, ami már  $n = 10$  esetén is igen időigényes, vagyis általánosan ez az út nem járható. Mi azonban végeztünk számításokat annak érdekében, hogy megtudjuk, érdemes-e ezzel a problémával behatóbban foglalkozni.

A Burr-Pareto-Logisztikus eloszlás a teszt gyengéjének bizonyult, itt már  $n = 8, 9$  méretű minták esetén is jól megfigyelhetőek a próbák erejében levő különbségek. Ezért ezen eloszlásra, valamint a  $\nu = 3$  szabadságfokú  $t$ -eloszlásra végeztünk számításokat, eredményeinket a 3. és 4. táblázatban mutatjuk be.

3. táblázat. Burr-Pareto-Logisztikus és  $t$ -eloszlás,  $m = 2$ ,  $n = 8$

	10%	5%	1%
$F$	0.4048	0.2846	0.1242
$L$	0.3076	0.2222	0.0956
$N$	0.4166	0.3012	0.1346
$T$	0.3506	0.2328	0.0818
$W$	0.2696	0.1750	0.0654
$W+$	0.3234	0.2056	0.0584
$H$	0.4876	0.3518	0.1514

	10%	5%	1%
$F$	0.2190	0.1530	0.0650
$L$	0.1874	0.1198	0.0484
$N$	0.2054	0.1332	0.0526
$T$	0.2176	0.1492	0.0678
$W$	0.3566	0.2600	0.1136
$W+$	0.2946	0.2170	0.0846
$H$	0.2790	0.2016	0.0998

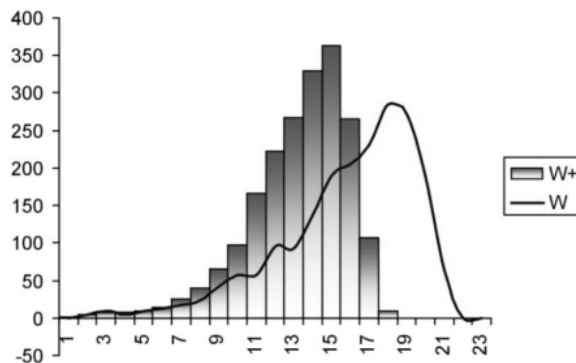
4. táblázat. Burr-Pareto-Logisztikus és  $t$ -eloszlás,  $m = 2$ ,  $n = 9$

	10%	5%	1%
$F$	0.4614	0.3304	0.1538
$L$	0.3528	0.2570	0.1192
$N$	0.4740	0.3526	0.1696
$T$	0.4052	0.2664	0.1018
$W$	0.2750	0.1816	0.0668
$W+$	0.3658	0.2312	0.0632
$H$	0.5446	0.4082	0.1862

	10%	5%	1%
$F$	0.2460	0.1782	0.0900
$L$	0.2118	0.1456	0.0638
$N$	0.2340	0.1582	0.0686
$T$	0.2442	0.1760	0.0872
$W$	0.3642	0.2730	0.1286
$W+$	0.3174	0.2300	0.0928
$H$	0.3004	0.2268	0.1194

A  $W$  sorok jelentik az eredeti implementációt, a  $W+$  sorok pedig a valós minimumban számított tesztre vonatkozó vizsgálatokat. Már ilyen kis mintáknál is jól kivehető, hogy a kívánt javulás bekövetkezett. Habár a  $t$ -eloszlás esetében a próbák ereje némileg csökkent, még így is igen jelentősen felülmúlja a többi. A  $H$ -tesztnél is magasabb marad, ami a második legerősebb volt a  $t$ -eloszlás ellenében. A 10. grafikonon a két teszt empirikus eloszlását ábrázoltuk az  $m = 2$ ,  $n = 8$  esetre,  $r = 2000$  ismételt mintavételezés esetén. A  $W+$  értékek, mint ahogyan várható volt, többnyire alacsonyabbak, mint a  $W$  értékek, ezért az eloszlás balra dől. A különbség látványos, ami azt jelenti, hogy a  $W$ -teszt kiszámítására adott eljárással valójában igen sokat tévedünk.

A próbaerő vizsgálatok alapján úgy látjuk, mindenképpen érdemes a  $W+$



10. ábra. A  $W$  és  $W+$  tesztek empirikus eloszlása

teszttel behatóbban foglalkozni. Mivel általában dominálja a többi tesztet, ezért alkalmas omnibusz tesztnek.

#### 5.4. Következtetések

A Malkovich-Affifi-féle tesztre vonatkozó vizsgálatok igen érdekes eredményekkel szolgáltak. A tesztnek voltak nyilvánvaló gyengéi, illetve más esetekben kimagaslott a többi teszt közül. Az elliptikus eloszlások a teszt egyértelmű erősségei, míg a  $\chi^2$  és Burr-Pareto-Logisztikus eloszlások esetében szerényebb teljesítményt nyújtott a többi alternatívával szemben. Az is érdekes eredmény, hogy a  $H$ -teszt, ami a peremekre vonatkozó információkat kombinálja, sok esetben vetekszik a  $W$ -teszttel. A peremek vizsgálata tehát, mint első lépés, mindenképpen célravezető a többdimenziós normalitás tesztelésében.

A  $W$ -teszt kiszámítására szolgáló eljárást korrigáltuk, mivel a tényleges megvalósítás nem felel meg az eredeti elképzelésnek. A

$$\min_c W(c)$$

teszt eloszlása más, mint a megvalósított eljárás, ami a teszt erejében is tükröződik. Egyelőre azonban nem tudunk általánosan használható módszert adni a valós  $W+$  teszt kiszámítására.

Pénzügyi alkalmazásokban, mint általában a közgazdaságtan egyéb területein, kezdetben a kézenfekvő többdimenziós normalitás hipotézisével éltek. Az újabb keletű kutatások azonban egyre inkább fordulnak egyéb eloszlások felé, mivel a normális eloszlással való közelítések hibái igen magasak. A normális eloszlás kedvező tulajdonságai közül valójában elégséges az eloszlás *elliptikus* tulajdonságaira építeni (*Embrechts et.al. [1999], Fang et.al. [1990]*), a pénzügytan szempontjából ezért az ún. *körkörös* (spherical) illetve *elliptikus* (elliptical) eloszlások családja, nem pedig maga a normális

eloszlás játszik központi szerepet. Ebbe a családba tartoznak — többek között — a többdimenziós Student  $t$ - és kevert normális eloszlások is, melyek a normális eloszlásnál fontosabb szerephez jutnak. Láttuk, hogy a többdimenziós Shapiro–Wilk teszt  $F$ ,  $L$ ,  $N$  és  $T$ , valamint  $H$ -változata gyengének bizonyul elliptikus  $H_1$  alternatívákra. Ez abból fakad, hogy a normális eloszlásnak egy olyan tulajdonságára építenek (linearitás, lineáris korreláció), mely az elliptikus eloszlásoknak is sajátja. Ezzel szemben a Malkovich és Afifi által javasolt változat érzékenynek mutatkozik elliptikus alternatívákra is, különösen pedig azokra, melyekben gyakoriak a kiugró értékek. Ezen eloszláscsaládon belül pedig érdeklődésre számíthat egy olyan teszt, mely speciálisan alkalmas elliptikus eloszlások — például a normális és Student  $t$ -eloszlások — szétválasztására.

Az ökonometriában illetve a statisztikában is fontos szerephez jut a  $t$ -modellek alkalmazása olyan adatmodellek esetében, melyeknél a hibatag eloszlása hosszan elnyúló, gyakoriak a kiugró elemek (outlier-ek). Ezen modellek alkalmazása valójában igen régre nyúlik vissza (*Jeffreys, [1939]*), és a területen napjainkban is aktív kutatások folynak.

A formálisan megadott eloszlásokkal végzett vizsgálatok mellett nyilván azonos fontosságú, hogy valós adatsorok esetében milyen teljesítményt nyújtanak az egyes tesztek. Két alkalmazást fogunk bemutatni, majd az alapadatokon bootstrap technikával kismintás próbákat fogunk végezni. Az első alkalmazásnál a peremekre vonatkozó normalitás intuitíve sejthető, melyet a formális tesztek is alátámasztanak. Az együttes normalitás tesztelését tehát ettől érzékenyebb módszerekkel is el kell végezni. A második alkalmazásnál viszont a perem információ lesz a döntő a normalitás hipotézisének elvetésében, hiszen mint látni fogjuk, teljesen különböző peremek mellett némelyik teszt el fogja a normalitást fogadni.

## 6. Alkalmazások

Gyakorlati példaként két pénzügyi alkalmazást választottunk. A többdimenziós normalitás hipotézise egy igen szigorú követelmény, és ebből a szempontból a pénzügy egy kritikus alkalmazási terület. A normalitás hiányának figyelmen kívül hagyása és a helytelen modellválasztás következményei igen súlyosak lehetnek. A legegyszerűbb arra az esetre gondolni, hogy befektetéseink várható kockázatát alulbecsüljük, melynek anyagi következményei beláthatatlanok. A pénzügyi piacon egy-két százalékos tévedés a várható veszteségekben millió forintos abszolút veszteségeket eredményezhet.

Emellett nem mellékes, hogy a pénzügyi idősorok vektorai felelnek meg legkevésbé az együttes normalitás követelményének más közgazdasági területekhez hasonlítva. Ez is az oka annak, hogy általánosabb, nemlineáris modellek fejlődését és alkalmazását ez a terület rendkívüli mértékben inspirálja. A többdimenziós normális eloszlás mellett fontosabb szerephez jutnak az

elliptikus eloszlások, ezen belül a  $t$ - illetve kevert normális eloszlású modellek igen kedveltek. A kopulák használatának elterjedésével ugyanakkor kiléphetünk az elliptikus eloszlások világából, és az elliptikus modelleknél rugalmasabb modelleket építhetünk.

Az első alpontban egy olyan alkalmazást mutatunk be, melynél két változó mindegyikére vonatkozóan az egydimenziós normalitás feltevése megállja a helyét, viszont az alkalmazott sokváltozós módszer az együttes normalitás teljesülését is megköveteli. Ez jó példa alkalmazás annak megállapítására, hogy mely tesztek érzékenyek a többdimenziós karakterisztikákra, és melyek azok, amelyek a perem információkra építenek. Itt azt próbáljuk meg illusztrálni szimulációs vizsgálattal, hogy a normalitástól való eltérés milyen mértékű torzítást visz az opcióárak várható értékére vonatkozó becslésekbe.

A második alkalmazás a piaci *kockázatkezelés* területére kalauzol el bennünket, ahol a normális eloszlásra épülő modellek kora már lejárt. A részvényhozamok együttes eloszlása illetve a részvényekből összeállítható portfóliók hozamának eloszlása jellemzően csúcsos, hosszan elnyúló. Példákon keresztül mutatjuk meg, hogy a  $t$ -modellek használata megfelelőbbnek bizonyul, mint a gaussi modelleké. A 5.2. pontból azt sejtjük, hogy az ilyen típusú eloszlásokra az általunk vizsgált  $W$ -teszt nagyobb érzékenységet mutat, mint a többi variáns. Célszerű megvizsgálni, hogy valós adatsorok esetében a dolog hogyan áll.

## 6.1. Opcióárazás numerikus módszerekkel

A tőzsde napjaink nagy érdeklődést kiváltó gazdaságelméleti területe. Az a felismerés, hogy a különböző értékpapírok árfolyamainak mozgása jól leírható sztochasztikus folyamattal, megnyitotta az utat a tőzsde, illetve a különböző értékpapírok és származékaik árfolyamainak matematikai modellezése irányába. Ehhez jelentős mértékben az a tény is hozzájárult, hogy az elméleti fizikai kutatások már foglalkoztak hasonló modellekkel, s így készen kínálták bonyolultabb differenciál egyenletek megoldásait; igaz, más mögöttes tartalommal. Különösen nagy figyelmet kaptak az opcióárazására vonatkozó modellek.

Ezen alpont jórészt a *Benedek [1999]* tanulmányra támaszkodik, melyben a szerző a *Black-Scholes*-féle formulát és a kapcsolódó (sztochasztikus) modellt mutatja be. Bizonyos ideális feltételek mellett az említett modellnek létezik analitikus megoldása, az opcióár egzakt módon megadható. A szerző a modell ideális világát tágítva feloldja az egyik alapfeltevést, és pedig a zérus tranzakciós költségekre vonatkozó kényszer feltételt. A feltétel feloldásával azonban az analitikus megoldásról le kell mondanunk, a *Black-Scholes* formula többé nem használható. Az általánosított esetben *diszkrét szimuláció* alkalmazásával a feladatot numerikusan kell megoldanunk. Egy valósághoz közelebb álló modellt készítünk, melyben az opció ára — mint látjuk majd — többé nem egy fix érték, hanem egy valószínűségeloszlás. A sztochaszti-

kus modellből számítógéppel generáljuk az opcióár ismeretlen  $G$  eloszlását, illetve abból egy véges mintát, majd pedig a kérdéses sokasági paramétert, vagyis az opcióár várható értékét ebből a mintából becsüljük.

Habár a statisztikai becslés csak közelítése (becslése) az elméleti értéknek, a mintaméret növelésével becslésünk varianciája egyre kisebb lesz, vagyis egyre biztosabban kerül az elméleti érték közelébe. A minta méretének növelésével azonban könnyen erőforráskorlátokba ütközhetünk, mivel egy szimulációs modell kiértékelése általában magas műveletigényű. Ahhoz, hogy a meglévő számítási kapacitáson belül maradjunk, de becslésünk varianciáját csökkentsük, illetve adott esetben hatásosságát növeljük, variancia csökkentő módszereket szükséges alkalmaznunk. A legkézenfekvőbb megoldás az, hogy nem független azonos eloszlású, vagyis FAE mintát generálunk a modellből, hanem *korrelált mintát*. Mint tudjuk, a FAE mintát könnyebb kezelhetősége miatt használjuk, azonban ennek a mintának a legkisebb az információtartalma. A várható értéknek az egyszerű mintaátlaggal való becslése azonban akkor is torzítlan marad, ha a mintaelemek nem függetlenül lettek mintavételezve, így tehát kézenfekvő, hogy a FAE mintából számított egyszerű átlag nem hatásos becslés.

Egy másik eljárás arra a technikai feltételre támaszkodik, hogy minden szimulációs modellben generálunk olyan véletlen változókat akár inputként, akár valamilyen közbenső segédváltozóként, melyeknek az eloszlása illetve sokasági paraméterei ismertek. Az opcióárazásos modell esetében például magát a piacot kell szimulálnunk általunk megválasztott (vagy becsült) paraméterekkel, vagyis ismerjük a modell inputját képező eloszlást. Emellett a modell inputja nyilván valamilyen szinten korrelál az output változóval is (hiszen éppen ezt az összefüggést vizsgáljuk), melynek eloszlását jellemezni kívánjuk, esetünkben a várható értékkel. Ezt a függőséget, továbbá azt a tényt kihasználva, hogy az input vagy egyéb közbenső segédváltozó eloszlását ismerjük, a vizsgált output változó ingadozását csillapíthatjuk. Ez az ún. *kontroll változós módszer*, ahol is a kontroll az ismert eloszlású segédváltozó vagy változó vektor.<sup>10</sup> Egy sokváltozós statisztikai módszerről van tehát szó, mely az

$$\text{Var}[y \mid \mathbf{c}] \leq \text{Var}[y] \quad (y, \mathbf{c}) \sim G_{k+1}(\hat{\theta}_y, \theta_c) \quad (12)$$

összefüggésre épít (Wolff [1989]), ahol is a  $\mathbf{c}$  kontroll változó  $k$ -dimenziós eloszlásának típusa és paraméterei ismertek, az  $y$  eredmény változó paraméterei becsülendők, eloszlásának típusa ismert vagy ismeretlen. Ebből következően a  $G_{k+1}$  együttes eloszlás nem ismert. Azt is tudjuk ugyanakkor, hogy

$$E[y \mid \mathbf{c}] = E[E[y \mid \mathbf{c} = \mathbf{C}]] = \mu_y \quad (13)$$

---

<sup>10</sup>A módszerről Nelson [1990] ad jó áttekintést.



a kétszeres várható érték tételből (*Bickel-Doksum [1977]*). Ahhoz azonban, hogy a (13) feltételes várható értéket becsülni tudjuk, szükséges a  $G$  eloszlásra vonatkozóan hipotézissel élnünk. Mint bemutattuk a 1.1. pontban, a (13) elsőfajú regresszió lineáris, ha  $G$   $k + 1$  dimenziós normális eloszlást követ.

Mielőtt azonban további részletekbe mennénk, bemutatjuk magát az opcióár modellt.

### 6.1.1. A Black-Scholes-féle opcióárazás és az „ideális” piaci feltevések

Vegyünk egy darab valamilyen részvényre vonatkozó vételi opciót (*call option*). A vételi opció valamilyen eszköz, esetünkben részvény vételére való jogot jelent egy meghatározott időn, a lejárat időn belül. Az *opció lehívása* azt jelenti, hogy az adott eszközt egy előre meghatározott áron, az ún. kötési árfolyamon a tulajdonosa megvásárolja. Az *amerikai típusú* opciónál a lejárat időn belül bármikor, az *európai típusú* opciónál csak a lejárat időpontban lehet az opciót lehívni. A lejáratig tartó idő az opció *futamideje*. Az általunk vizsgált modellben a vételi opció európai típusú.

Mi az értéke egy - a  $T$  lejárat idő előtti -  $t$  időpontban az európai vételi opciónak? Tegyük fel, hogy a részvény árfolyama, melyre az opció vonatkozik, a  $t$  időpontban  $S_t$ . Ha a kötési árfolyam  $E$ , akkor a  $t$  időpontban a jelenértéke  $E^{(t)} = Ee^{-r(T-t)}$ . Ekkor ha  $S_t > E^{(t)}$ , akkor az opciót le fogják hívni, így az opció értéke  $S_t - E^{(t)}$ , míg ha  $S_t < E^{(t)}$ , akkor az opciót nem fogják lehívni, vagyis értéke nulla. Formálisan

$$C(S_t, t) = \max \{S_t - Ee^{-r(T-t)}, 0\} = e^{-r(T-t)} \max \{S_T - E, 0\} \quad (14)$$

ahol  $C(\cdot)$  az opció értékét jelöli a  $t$  időpontban, ha  $S_t$  a  $t$  időpontbeli árfolyam. Látható, hogy ha a  $t$  időpont a lejáratához közel esik, akkor az opció értéke a részvényárfolyam mínusz a kötési árfolyam. Ha viszont a lejárat nagyon távoli, akkor kötési árfolyam jelenértéke elhanyagolható a részvényárfolyamhoz képest, így az opció értéke megegyezik a részvény árával.

Ez lenne a helyzet, ha a részvényárfolyam determinisztikus folyamatot követne, vagyis a (14) képletben  $S_T$  értékét pontosan meg tudnánk határozni. Csakhogy a részvény árfolyama sztochasztikus folyamatot követ, így az előbbi egyszerű jelenérték-számítás nem működik a valós esetben.

Nézzük meg, milyen sztochasztikus folyamatok alkalmasak az árfolyam mozgásának leírására? Ehelyütt előre bocsátjuk, hogy mivel a Black-Scholes-féle modell folytonos, mi pedig a következő pontban szükségszerűen diszkrét modellt fogunk építeni, ezért a sztochasztikus folyamatok ismertetésénél bemutatjuk mind a folytonos, mind pedig a diszkrét változatokat.

A *fehér zajt* alap statisztikai tanulmányunkból már ismerjük.

**68. Definíció.** Az  $S_t$  sztochasztikus folyamat fehér zaj, ha

$$\begin{aligned} S_t &= \varepsilon_t \quad \varepsilon_t \sim N(0, \sigma), \quad t = 0, 1, \dots, T && \text{diszkrét esetben} \\ S(t) &= \varepsilon(t) \quad \varepsilon(t) \sim N(0, \sigma), \quad t \in [0, T] && \text{folytonos esetben} \end{aligned}$$

A fehér zaj folyamatra diszkrét és folytonos esetben is igaz, hogy  $E[S_t] = 0$ ,  $Var[S_t] = \sigma^2$  és  $Cov[S_t, S_k] = 0$  ( $t \neq k$ ), azaz várható értéke minden időpontban zérus, szórása  $\sigma$  és két különböző időpontban felvett érték egymástól független. A *bolyongási folyamat* a fehér zajból származik:

**69. Definíció.** Az  $S_t$  sztochasztikus folyamat véletlen bolyongás, ha

$$\begin{aligned} S_t &= S_{t-1} + \varepsilon_t, \quad t = 0, 1, \dots, T && \text{diszkrét esetben} \\ dS(t) &= \varepsilon(t) \quad t \in [0, T] && \text{folytonos esetben} \end{aligned}$$

ahol  $\varepsilon_t$  (illetve  $\varepsilon(t)$ ) fehér zaj.

A bolyongási folyamatnak tehát a differenciája fehér zaj. Diszkrét esetben igazak a következő összefüggések:

$$\begin{aligned} E[S_t] &= E[S_{t-1} + \varepsilon_t] = \dots = E\left[S_0 + \sum_{k=1}^t \varepsilon_k\right] = E[S_0] = S_0 \\ Var[S_t] &= Var\left[S_0 + \sum_{k=1}^t \varepsilon_k\right] = Var[S_0] + \sum_{k=1}^t Var[\varepsilon_k] = t\sigma \\ Cov[S_t, S_{t-s}] &= E\left[\left(S_0 + \sum_{k=1}^t \varepsilon_k\right)\left(S_0 + \sum_{k=1}^{t-s} \varepsilon_k\right)\right] = E\left[\sum_{k=1}^{t-s} \varepsilon_k^2\right] = (t-s)\sigma^2 \end{aligned}$$

Az előbbiek folytonos esetre is igazak, bár bizonyításuk nem ennyire kézenfekvő, mivel az  $\varepsilon(t)$ -nek minden pontjában szakadása van, így például az  $\int_{k=0}^t \varepsilon(k) dk$  integrál nem létezik.

Az előbbiek szerint tehát  $S_0$ -tól indulva a fehér zaj folyamat várható értéke a  $t$  időpontban zérus, míg a bolyongási folyamaté  $S_0$ . Ha azonban a várható értéket becsülni szeretnénk, akkor a becslés távolabbi időpontokra utóbbi esetében egyre bizonytalanabb, hiszen a variancia  $t$ -ben növekszik. Ez megfelel a részvényt piac tulajdonságainak. Ha a  $t-1$  időpontban szeretnénk megbecsülni a  $t$  időpontbeli értéket, akkor ez fehér zaj esetében megint csak zérus, míg a bolyongási folyamat esetében  $S_{t-1}$ . Ezek alapján azt mondhatjuk, hogy a részvényárfolyamok is valamifajta bolyongásszerű mozgást végeznek. Minél hosszabb távra becsüljük előre részvényünk árfolyamát, annál kevesbé vagyunk biztosak abban, hogy a becsült árfolyam majd a valós árfolyam közelében lesz.

Mivel az árfolyammozgások általában trendet is tartalmaznak, ezért modelleként célszerűbb a trenddel bővített

$$\begin{aligned} S_t &= S_{t-1} + \mu + \varepsilon_t, \quad t = 0, 1, \dots, T && \text{diszkrét esetben} \\ dS(t) &= \mu + \varepsilon(t) \quad t \in [0, T] && \text{folytonos esetben} \end{aligned}$$

alakokat felhasználni. Egy speciális bolyongásos folyamat a Wiener-folyamat:

**70. Definíció.** A  $z_t$  sztochasztikus folyamat Wiener-folyamat, ha

$$\begin{aligned} \Delta z_t &= \varepsilon_t \sqrt{\Delta t} \quad \varepsilon_t \sim N(0, 1) && \text{diszkrét esetben} \\ dz(t) &= \varepsilon(t) \sqrt{dt} \quad \varepsilon(t) \sim N(0, 1) && \text{folytonos esetben} \end{aligned}$$

vagyis a  $z(t)$  ( $z_t$ ) változása normális eloszlású zérus várható értékkel és  $dt$  ( $\Delta t$ ) varianciával.

Hasonlóan a véletlen bolyongáshoz, könnyen levezethető, hogy a  $z_t$  folyamat várható értéke  $E[z_t] = z_0$ , varianciája pedig az eltelt idő, vagyis  $Var[z_t] = t$ . A Wiener-folyamat általánosítható olyan módon, hogy kibővítjük trenddel, valamint alkalmassá tesszük különböző kockázatú értékpapírok modellezésére. Legyen tehát az általánosított modell

$$\Delta S_t = a\Delta t + b\Delta z_t, \text{ azaz } S_t = S_{t-\Delta t} + a\Delta t + b\varepsilon_t \sqrt{\Delta t} \quad \text{diszkrét esetben} \quad (15)$$

$$dS(t) = a dt + b dz(t) \quad \text{folytonos esetben}$$

Az  $a$  együttható a trend - vagy drift -, a  $b$  paraméter pedig a variancia - vagy volatilitás - meghatározója. Bizonyítható, hogy

$$\begin{aligned} E[\Delta S_t] &= a\Delta t \\ Var[\Delta S_t] &= b^2 \Delta t \end{aligned}$$

valamint

$$S_t \sim N(S_0 + at, b^2 t)$$

vagyis az általánosított Wiener-folyamatot követő részvény árfolyama minden időpillanatban a fenti paraméterekkel rendelkező normális eloszlású valószínűségi változó.

Ha feltesszük, hogy  $S_0$  a részvény induló árfolyamát jelöli, és a drift 12, akkor egy időszak múlva  $E[\Delta S_t] = a\Delta t = 12$ , vagyis  $S_t = S_0 + 12$ . Ez pedig azt jelenti, hogy  $S_0 = 100$  esetén  $S_t = 112$ ,  $S_0 = 1000$  esetén  $S_t = 1012$ . Mi azt szeretnénk, ha mindkét esetben 12%-os lenne a változás, ezért célravezetőbb  $S_t$ -vel az árfolyam logaritmusát jelölni. Ekkor ui.  $\Delta \log(S) = \log(S_t/S_{t-1})$  és  $d \log(S) = dS/S$ , vagyis a (15) képletekben a bal oldalon a változás helyett a százalékos változást szerepeltetjük. Az árfolyamok logaritmusában a drift paraméter által indukált additív változás az árfolyamokban már multiplikatív, vagyis százalékos változást eredményez.

**71. Megjegyzés.** Mivel minden  $t$  időpillanatra vonatkozóan  $\log(S_t)$  követ normális eloszlást, ezért  $S_t$  lognormális eloszlást követ, mégpedig

$$S_t \sim LN\left(S_0 e^{at}, S_0^2 e^{2at} (e^{b^2 t} - 1)\right)$$

paraméterekkel.

A lognormális eloszlás ugyanakkor alkalmasabb is részvényárfolyamok modellezésére, mivel nem vesz fel negatív értéket, és nagyobb valószínűséget enged kiugróan magas árfolyamoknak, mint a normális eloszlás.

Látható tehát, hogy a normális eloszlás feltevése itt sem állja meg a helyét. Az extrém piaci események modellezésére ugyanis ez az eloszlás nem alkalmas. Az árfolyamok esetében a jobbra hosszán elnyúló eloszlások, mint például a lognormális eloszlás sokkal realisabb feltevés<sup>11</sup>. Ebben az esetben azonban egy olyan eloszlást használunk, mely visszavezethető normális eloszlásra (ui. az árfolyamok logaritmusára tesszük fel a normalitást).

**72. Megjegyzés.** A teljesség kedvéért megemlíjtük, hogy az árfolyamok logaritmusában felírt általánosított Wiener-folyamat az árfolyamokban egy másik nevezetes sztochasztikus folyamat, ui.

$$d \log(S) = dS/S = adt + bdz \implies dS = aSdt + bSdz$$

amit geometriai Brown-mozgásnak nevezünk.

A sztochasztikus folyamatokba való rövid bevezető után nézzük meg magát a Black-Scholes modellt! Tételizzük fel tehát, hogy egy részvényre vonatkozó vételi opciót vásárolunk, a részvényárfolyam logaritmus általánosított Wiener-folyamatot követ  $\mu$  és  $\sigma$  paraméterekkel, vagyis

$$\begin{aligned} d \log(S) &= \mu dt + \sigma dz \text{ illetve a 72. megjegyzés alapján} \\ dS &= \mu S dt + \sigma S dz \end{aligned} \quad (\text{F1})$$

A vételi opció árfolyamát egy

$$C = F(S, t)$$

alakú — egyelőre ismeretlen — függvény határozza meg. Feltesszük továbbá, hogy a piaci kamatláb, mely kockázatmentes befektetést biztosít, az opció futamideje alatt konstans, vagyis

$$r = \text{konstans a futamidő alatt} \quad (\text{F2})$$

---

<sup>11</sup>Mint később látni fogjuk, hozamok esetében pozitív és negatív extrém események bekövetkezését kell alkalmas eloszlással leírunk.

További feltevés, hogy a részvény a futamidő alatt nem fizet osztalékot, tehát

$$r_S = 0 \quad (\text{F3})$$

Nincsenek tranzakciós költségek. Lehetőség van ún. *short selling*-re, azaz eladhatunk úgy egy részvényt valakinek, hogy az nincs a birtokunkban, csak megegyezés szerint helyt kell állnunk érte valamikor a jövőben. A feltételezés szerint a short selling-nek nincsenek többletköltségei. Nincs továbbá költsége a kölcsönvételnek sem, azaz lehetőségünk van kockázatmentes kamatláb mellett kölcsönt felvenni. Minden időpillanatban — folytonosan — lehetőség van kereskedésre. Legyen tehát

$$TC = 0 \quad (\text{F4})$$

$$\text{európai típusú opcióról van szó} \quad (\text{F5})$$

vagyis csak a lejáratkor hívható le, és

$$\text{nincs lehetőség arbitrázsra.} \quad (\text{F6})$$

A Black-Scholes képlet tehát elvileg csak olyan ideális körülmények közt használható, amire a világon sehol sincs példa<sup>12</sup>. Ennek ellenére a formulát mégis előszeretettel alkalmazzák, és beépítik sok kockázatkezelő szoftverbe.

Tekintsünk egy olyan portfóliót, ahol eladunk egy darab vételi opciót és vásárolunk  $F_S$  darab részvényt és  $F_S = \partial F / \partial S$ . Ez valami olyasmit jelent, hogy amint megváltozik az opció ára, azonnal módosítjuk portfóliónkat. Ekkor a portfóliónk értéke

$$\Pi = -F(S, t) + F_S S$$

a portfólió értékének változása pedig

$$\begin{aligned} d\Pi &= -dF(S, t) + F_S dS = -F_S dS - F_t dt - \frac{1}{2} F_{SS} (dS)^2 + F_S dS \\ &= -F_t dt - \frac{1}{2} F_{SS} \sigma^2 S^2 dt \end{aligned} \quad (16)$$

A levezetésben felhasználtuk az ún. Ito-lemmát, ennek leírását lásd *Black-Scholes [1973]* és *Hull [1993]* művekben. Ha megvizsgáljuk a kapott összefüggést, azt tapasztaljuk, hogy a sztochasztikus változót tartalmazó tag  $(dS)$  kiesett, vagyis portfóliónk értékének változása nem függ a véletlentől. Ezek szerint portfóliónk mindaddig kockázatmentes marad, ameddig a részvényárfolyam változására azonnal reagálva kiegészítjük portfóliónk értékét.

<sup>12</sup>A modell általánosításai megtalálhatóak a következő művekben: *Cox-Ross [1976]*, *Hull-White[1987]*, *Cox et al. [1979]*.

Ezt a szakirodalom dinamikus fedezésnek (*dynamic hedging*) hívja. Mivel portfóliónkat ilyen stratégiával kockázatmentesen tudjuk tartani, ezért a portfólió értékének növekménye meg kell hogy egyezzen a portfólió értékének kockázatmentes kamattal számított növekedésével. Ellenkező esetben arbitrázsra lenne lehetőség, vagyis

$$d\Pi = \Pi r dt \quad (17)$$

Ha a (17)-be behelyettesítjük a (16) összefüggést, kiesik a  $dt$  tag, és visszamarad egy differenciál egyenlet

$$F_t + rSF_S + \frac{1}{2}F_{SS}\sigma^2S^2 = rF(S, t)$$

ahol az ismeretlen az opcióár függvénye ( $F(S, t)$ ). Ahhoz, hogy az egyenletnek egyetlen megoldása legyen, szükség van egy kezdeti feltételre. Kezdeti feltételként a (14) összefüggést használjuk  $t = T$  helyettesítés mellett. Ekkor a feladat átalakítható egy olyan parciális differenciál egyenletté, mely a fizikában ismert hővezetés egyenlete és megoldása ismert (*Churchill [1963]*):

$$F(S, t) = S\Phi(d_1) - Ee^{-r(T-t)}\Phi(d_2) \quad (18)$$

ahol

$$d_1 = \frac{\ln(S/E) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

$$d_2 = \frac{\ln(S/E) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} = d_1 - \sigma\sqrt{T-t}$$

és  $\Phi(\cdot)$  a standard normális eloszlás eloszlásfüggvénye.

### 6.1.2. Opcióárazás tranzakciós költségek mellett

Ebben az alpontban bemutatjuk, hogy sikerült az opció árát meghatározni olyan esetben, amikor analitikus képlet nem áll rendelkezésre. Az (F4) feltételt fogjuk feloldani, azaz bevezetjük a modellbe a tranzakciós költségeket. Mint láttuk, optimális esetben a piaci szereplő a dinamikus hedge-lési stratégiát választja, vagyis a részvény árfolyamának változására azonnal reagál, mégpedig „folytonosan“. Ezt megteheti, hiszen tetszőleges számú, kontinuum számosságú tranzakciót lebonyolíthat költségmentesen és bármikor további pénzforrásokhoz juthat.

A tranzakciós költségek bevezetésével azonban a helyzet döntően megváltozik. A kiegészítések magas száma ugyanis magas tranzakciós költségekkel jár együtt. A piaci szereplő megteheti, hogy csökkentvén tranzakciós költségeit, kevesebbszer egészíti ki portfólióját pl. megérzéseire alapozva, de mint később látni fogjuk, így fokozott kockázatnak fogja magát kitenni.

Felépítünk egy olyan modellt, mely minden egyes kiegészítéshez  $TC$  nagyságú tranzakciós költséget rendel, valamint felírunk egy döntési szabályt arra vonatkozóan, hogy a piaci szereplő mikor, milyen események hatására egészíti ki a portfólióját. Ellentétben a Black-Scholes-féle modellel, mi egy diszkrét modellt építünk, ahol is a kontinuum idő tengelyt felváltjuk napokkal, hónapokkal, negyedévekkel stb. Az események tehát diszkrét léptékekben történnek, így a döntés alapja sem a pillanatnyi, hanem a napi záróárfolyam, illetve a megfelelő időintervallumonként (nap, hónap, negyedév stb.) annak megváltozása. A dinamikus hedge-lés ebben az esetben tehát nem folyamatos, hanem szintén diszkrét egységekben valósul meg, s így az idő „kellően finom” felosztása esetén a Black-Scholes-féle eredmény jó közelítését kell visszakapnunk.

**Bemenő adatok generálása.** Első lépcsőben elő kell állítanunk magát a piaci környezetet, vagyis a részvény idősorát. A (15) képletet  $S_t$  helyett  $\log(S_t)$ -re alkalmazva a

$$\begin{aligned}\log(S_{t+\Delta t}/S_t) &= (\mu - \sigma^2/2) \Delta t + \sigma \varepsilon \sqrt{\Delta t} \implies \\ S_{t+\Delta t} &= S_t e^{(\mu - \sigma^2/2) \Delta t + \sigma \varepsilon \sqrt{\Delta t}}\end{aligned}\tag{19}$$

rekurzív formulát kapjuk, ahol  $\varepsilon \sim N(0, 1)$ . Mivel a modell logaritmikusokra van felírva, ezért a drift illetve volatilitást meghatározó paramétereket is ennek megfelelően átváltottuk.<sup>13</sup> A paraméterek a következők:

$MU$  a Wiener-folyamat  $\mu$  paramétere

$SI$  a Wiener-folyamat  $\sigma$  paramétere

$S_0$  a részvény induló árfolyama

$N$  hány naponként generáljunk új részvényárfolyamot

$T$  időszak hossza

$EV$  egy év hány nap hosszú legyen

$SSIZE$  mintaméret

Például ha  $N = 1$ ,  $EV = 365$  és  $T = 30$ , akkor egy naptári hónap az időszak hossza, és  $\Delta t = N/EV = 1/365$ . Ekkor egy  $T$  hosszúságú idősort generálunk, melynek a kezdő értéke konstans, éspedig  $S_0$ . Az  $EV$  változóra megadhatunk hipotetikus számokat is, például  $EV = 730$  értéket, ekkor szimulálhatunk félnapos, illetve kellően nagy  $E$  esetén percenkénti árfolyam szcenáriókat is.

---

<sup>13</sup>Ha  $Y \sim N(\mu_Y, \sigma_Y^2)$ , akkor  $X = e^Y \sim LN(\mu_X, \sigma_X^2)$ , ahol  $\mu_X = e^{\mu_Y + \sigma_Y^2/2}$  és  $\sigma_X^2 = e^{2\mu_Y + \sigma_Y^2} (e^{\sigma_Y^2} - 1)$ .

Egy idősor, vagyis egy szcenárió a piaci folyamat egy lehetséges megvalósulása, amiből az opcióár egy lehetséges értékét határozhatjuk meg. Mivel azonban a folyamat nem determinisztikus, ezért az opcióár várható értékének meghatározásához több különböző szcenárió eredményét kell számításba venni, vagyis az opcióárra vonatkozóan egy mintát kell generálnunk. A *SSIZE* paraméterrel állítjuk be a szcenáriók számát, vagyis a minta méretét, melyből az opcióár várható értékét kívánjuk megbecsülni. Az idősorhoz szükséges kiinduló standard normális eloszlásból származó pszeudo-véletlen számokat a 4. pontban már ismertettük.

**A szimuláció.** Második lépésben definiáljuk a piaci szereplő viselkedését és a piaci mechanizmusokat. A paraméterek a következők:

$E$  az opció kötési árfolyama

$R$  a kockázatmentes kamatláb

$TC$  tranzakciós költség nagysága százalékosan<sup>14</sup>

A modell változói:

$P$  egyenleg

$S$  részvényárfolyam

$\delta$  az opció deltája (fedezeti aránya)

Első nap tudjuk, hogy a részvény árfolyama az indulóárfolyam. Elkészítjük az első napi portfóliónkat: eladunk egy darab vételi opciót és vásárolunk  $\delta$  darab részvényt, ahol a (18) formulát alapul véve

$$\delta \equiv F_S = \Phi(d_1)$$

Tegyük fel, hogy az opcióért semmit sem kapunk, míg a részvények vásárlását  $R$  kamat mellett kölcsönből finanszírozzuk. Pénzünk így az első nap

$$P_0 = -S_0\delta_0(1 + TC)$$

Látható, hogy a tranzakciós költséget a kereskedés összértékével arányosan adjuk meg (nincs fix minimumköltség). A további napokban mindig ugyanaz történik, egészen az utolsó napig. Először megfizetjük pénzünkre (kölcsön) a kamatot:

$$P_t = P_{t-N}e^{R \cdot N/EV}$$

---

<sup>14</sup>Azaz  $TC = 0.01$  esetén 100 forint értékű részvény eladásának/vételének tranzakciós költsége 1 forint.



Ezután vesszük a soron következő  $S_t$  részvényárfolyamot, segítségével kiszámoljuk az új  $\delta$ -át. A portfóliónkban lévő részvények számát erre az értékre kell beállítanunk, tehát vagy eladunk, vagy veszünk további részvényeket. Pénzünk a következőképpen változik:

$$dP_t \begin{cases} S_t (\delta_{t-N} - \delta_t) (1 - TC), & \text{ha } \delta_{t-N} \geq \delta_t \\ -S_t (\delta_t - \delta_{t-N}) (1 - TC) & \text{különben} \end{cases}$$

Az utolsó napon hasonlóan az előzőekhez megfizetjük a kamatokat, és beolvassuk az utolsó naphoz tartozó részvényárfolyamot. A portfóliónkat viszont már nem egészítjük ki, sőt eladjuk részvényeinket, és helytállunk az opciónál, azaz:

$$dP_T = S_T \delta_{t-N} (1 - TC) - \max\{S_T - E, 0\}$$

Végül az első napra diszkontáljuk a kapott értéket és vesszük a mínusz egy-szeresét:

$$C = -P_T e^{-R \cdot T / EV}$$

A kapott érték ( $C$ ) az opció árát adja meg a periódus elején, hiszen ha pont ennyiért adtuk volna el az opciót az első periódusban, akkor az utolsó periódusban pénzünk nullával lenne egyenlő. Az opcióár, illetve opcióárfolyam tehát azt az értéket jelenti, mely az adott scenárió megvalósulása esetén zérus időszak végi egyenleget adna, ha a piaci szereplő optimális stratégia alapján döntene a teljes időszakban. Az előbbieken definiált modellt *BSTC* modellnek (Black-Scholes Transaction Costs) neveztük el.

Ahhoz, hogy az opcióár várható értékét becsülni tudjuk, a kísérletet többször (*SSIZE*-szor) meg kell ismételni. Az opcióárat első lépésben az egyszerű mintaátlaggal becsüljük. Az átlag mellett azonban szórás is számolunk, hiszen az opcióár most egy eloszlás. Bizonytalanság mellett fogjuk a döntésünket meghozni, ezért azt is fontos tudnunk, hogy döntésünk mekkora kockázattal jár, vagyis hogy a lognormális eloszlású opcióárnak mekkora a varianciája (illetve szórása).

A fenti modellt nagyon sok különböző paraméterre végigszámoltuk, és minden esetben az 5. táblázattal konzisztens eredményeket kaptunk, így csak a következő paraméter-beállítások eredményeit kívánjuk bemutatni:

$$MU = 0.12 \quad SI = 0.30 \quad S_0 = 100$$

$$E = 100 \quad R = 0.05 \quad SSIZE = 5000$$

Az opció futamideje 30 nap volt, így a Black-Scholes formula szerint az opció értéke: 3.6321.

Jól látható, hogy a modell zérus tranzakciós költségek mellett két tizedesjegy pontossággal megközelíti a Black-Scholes értéket. Azt is tapasztalhatjuk, hogy minél kisebbre választjuk a lépésközt, annál alacsonyabb a

5. táblázat. A BSTC modell eredményei

Hányszor egészítünk ki?	$TC = 0$		$TC = 0.01$	
	Várható érték	Szórás	Várható érték	Szórás
Naponta $5x$ (150)	3.6310	0.2414	8.6447	1.5247
Naponta $2x$ (60)	3.6309	0.3769	7.1760	1.0877
Naponta (30)	3.6333	0.5280	6.4315	0.9710
Kétnaponként (15)	3.6211	0.7327	5.5895	1.0325
Ötnaponként (6)	3.6240	1.1391	5.4094	1.3448
Csak első nap (1)	3.6207	2.5572	4.6957	2.5631
Soha (0)	3.8940	5.5273	3.8940	5.5273

szórás. Határértékben nyilván eltűnik a szórás, az opcióár eloszlása egyetlen ponttá zsugorodik, és ezt állítja a Black-Scholes levezetés is. Vagyis ha opcióval ügyletelünk, és adott esetben nincsenek (vagy nem kereskedés arányosak) a tranzakciós költségek, ne habozzunk olyan gyakran kiegészíteni portfóliónkat, amilyen gyakran csak lehet!

Zérus tranzakciós költségek mellett az opcióár várható értéke bármilyen lépésköz esetén a Black-Scholes formula által számított érték. A lépésköz az opcióár eloszlásának szórás paraméterét befolyásolja, a várható értékét nem. Abban az esetben azonban, ha vannak tranzakciós költségek, minél többször kereskedünk, annál magasabb értéket kapunk az opcióra. Határértékben az opció ára bármilyen pozitív tranzakciós költség esetén a végtelenbe tart. A kiegészítések növelésével azonban tranzakciós költségek esetén is csökkenthető az opció kockázatossága<sup>15</sup>. Ezek szerint, ha csökkenteni szeretnénk kockázatunkat (egy bizonyos fokig), gyakrabban kell kiegészítenünk, vállalva ezzel az esetleges kisebb nyereséget, míg ha kevésbé vagyunk érzékenyek a kockázatra, akkor egészítsünk ki ritkábban, vállalva azt, hogy a várt nagyobb hozam mellett esetleg nagyot bukunk. Az alacsonyabb kockázat tehát alacsonyabb várható nyereséggel (magasabb opcióárral) párosul és fordítva.<sup>16</sup> Összefoglalva, a *BSTC* modell konzisztens eredményt adott a Black-Scholes formulával és várakozásainknak megfelelő értékeket nyújtott a tranzakciós költségek bevezetésekor.

A továbbiakban a modell egy olyan, adaptív változatát mutatjuk be, melyben a lépésközt, vagyis a kiegészítések számát nem rögzítjük előre, hanem a szimuláció során a helyzetnek megfelelően alakítjuk. A képzeletbeli piaci szereplő az öt jellemző küszöbértékek alapján dönt arról, hogy változtat-e

<sup>15</sup>Egyre sűrűbb kiegészítés esetén azonban az opcióár-emelkedés mellett növekedni kezd a szórás. Ettől az értéktől kezdve a kiegészítések már biztosan nem hatékonyak.

<sup>16</sup>A várható értékre és a szórásra vonatkozóan egy vektor maximum problémával találkozunk, ahol a kétdimenziós céltér minden pontja efficiens. Erre a halmazra készíthetnénk egy hasznossági függvényt, amelynek maximalizálása megadja az optimális viselkedési stratégiát. Ekkor a különböző hasznosságfüggvények különböző befektetéstípusokat jellemeznének.

portfólióján, vagy sem.

**Optimális stratégia keresése, a  $K$ - $DH$  modell.** Az alapötlet az, hogy ha nem módosíthatjuk minden időpillanatban a portfóliónkat, akkor jobb lenne azokban az időpontokban módosítani, amikor arra a leginkább szükség van. Képzeljünk el egy tűréshatár sávot, és mondjuk azt, hogy abban az esetben, ha a megkívánt (számított) *delta érték* ( $\delta$ ) és a jelenlegi részvénymennység (egy korábbi delta) különbsége a sávon belül mozog, akkor nem változtatunk a portfólió összetételén, de ha a sávból kimozdul, akkor a megkívánt delta értékre korrigáljuk a részvények számát. Minél kisebb ez a tűréshatár, annál inkább közelítünk a  $BSTC$  modell minden időpontbeli kiegészítés eredményéhez, és minél szélesebb, annál inkább tartunk a kiegészítés nélküli eredményhez.

A szimulációs megvalósítás felépítésében egyezik az előző modellel, csak néhány feltételt kellett beiktatnunk. Megjegyezzük továbbá, hogy két  $K$ - $DH$  modellt készítettünk, az elsőben a kezdeti periódusban a sávnagyságtól függetlenül mindig kiegészítettünk ( $K$ - $DH1$ ), míg a másikban az első periódusra is vonatkozott a kiegészítési feltétel ( $K$ - $DH2$ ).

Maga a feltétel a következő:

$$vanKieg = \begin{cases} igaz & ha, |\delta_{aktuális} - \delta_t| \cdot S_t \geq S_0 \cdot TC \cdot K \\ hamis & különben \end{cases}$$

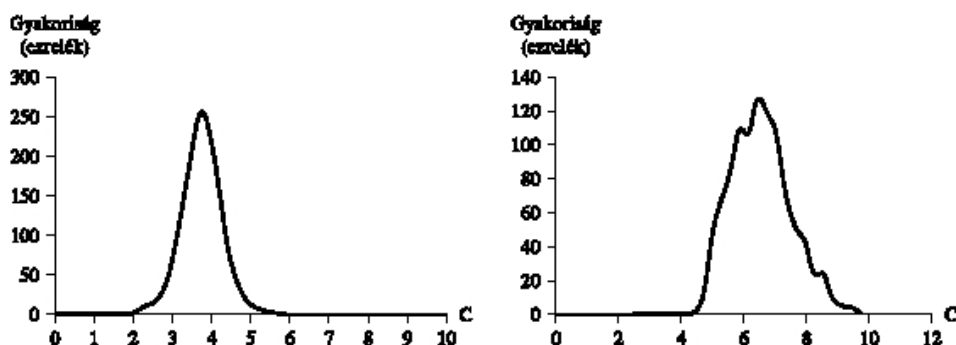
A jobb oldal a  $K$  paraméter függvényében egy sávot jelöl ki, ez a tűréshatár, magát a  $K$  paramétert a *tűréshatár sáv szélességének* neveztük el. A  $K$  paraméter a modell futtatása alatt egy rögzített érték, egy bizonyos preferenciájú — kockázat kerülő vagy kockázat kereső — piaci szereplőt testesít meg, és értéke a nem negatív számokon értelmezett.  $K = 0$  beállításnál a szereplő minden periódusban kiegészít, míg  $K = \infty$  határesetben soha sem egészít ki.

Legrészletesebben ezt a modellt vizsgáltuk, s ennek a modellnek a vizsgálata során derült fény néhány további kérdésre. Nézzük először a fenti paraméterbeállítások mellett, milyen eredményeket adott a modell ( $N = 1$ ,  $T = 30$ ,  $EV = 365$ ,  $TC = 0.01$ ) (6. táblázat). A 6. táblázatban jól látható, hogy minél inkább kockázatkerülő a befektető, annál kisebb  $K$  paraméterrel jellemezhető, és annál magasabb értéket kap az opció várható értékére. Az is látható, hogy ha a paramétert nagyon nagyra választja, akkor a futamidő alatt egyáltalán nem kereskedik a befektető, és így az opció árát nem terheli a tranzakciós költség, tehát az érték azonos az 5. táblázat megfelelő értékével. Miután elkészítettük a modellt, és meggyőződünk a modell helyességéről, az opcióár eloszlását vizsgáltuk.

A 11. ábra bal oldali grafikonja azt az esetet mutatja, amikor nulla tranzakciós költség mellett mindennap kiegészítettük a portfóliónkat. A várható érték nyilván a Black-Scholes megoldás, és a futamidő felosztásának

6. táblázat. A K-DH2 modell eredményei

$K$ paraméter	Várható érték	Szórás	Átlagos kiegészítések
0.0	6.4315	0.9710	30
4.0	6.2544	0.9886	16
13.0	5.7128	1.0662	6
48.0	5.0740	2.2766	2
999.0	3.8940	5.5273	0

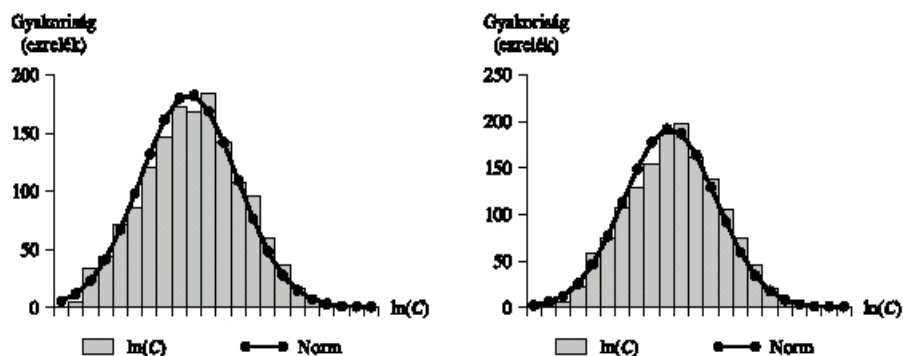


11. ábra.

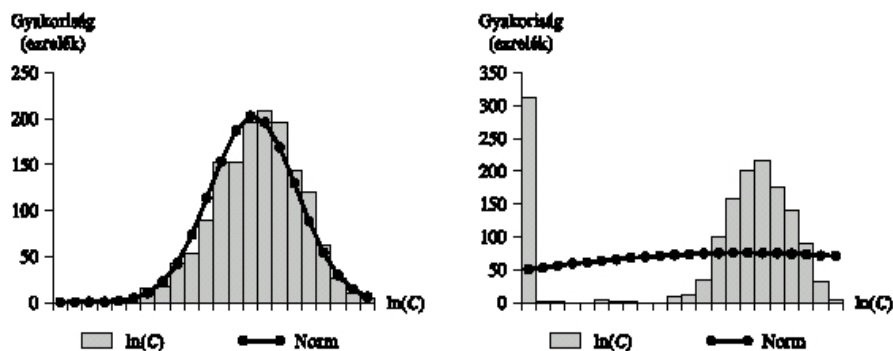
Opcióárak eloszlása a *BSTC* modellnél valamint a *K-DH* modellnél a  $K = 0$  esetben

finomításával a görbe egyre csúcsosabbá tehető (a szórás csökken), míg folytonos esetben ráhúzódik a Black-Scholes értékre (eltűnik a szórás). A 11. ábra jobb oldali grafikonja a tranzakciós költségekkel kibővített modell eredményét mutatja, ahol a  $K$  paramétert zérusnak vettük, s így ugyanúgy, mint a *BSTC* esetben, mindennap kiegészítettük a portfóliónkat. Jól látható, hogy a várható érték és a szórás megnövekedett, az eloszlás pedig log-normálisnak tűnik. Az előzőekben láttuk, hogy a  $K$  paraméter növelésére a várható érték csökken és a szórás növekszik.

A 12-13. ábrák azt is megmutatják, hogy  $K$  növelésével az eloszlás is drámai változáson megy keresztül. Lognormális illesztés helyett az opcióárak logaritmusára próbáltunk normális eloszlást illeszteni. Látható, hogy nem túl nagy  $K$  értékekre igen jó az illeszkedés, de ahogy nő  $K$ , úgy tolódik ki jobbra az eloszlás csúcsa, továbbá megfigyelhető, hogy egyre nagyobb az esélye az extrém kicsi értékeknek. Egy bizonyos paramétertől aztán szétrobban az eloszlás. Ekkor az történik, hogy bizonyos esetekben az opció értéke zérus. (Nagy  $K$ -ra ugyanis elképzelhető, hogy a szimuláció során egyszer sem történik kiegészítés, és az utolsó periódusban pedig nem kell helytállni az opciónál, mivel a részvény ára nem haladja meg a kötési



12. ábra. A  $K$ - $DH$  modell eredménye  $K = 0$ -ra és  $K = 4$ -re



13. ábra. A  $K$ - $DH$  modell eredménye  $K = 16$ -ra és  $K = 32$ -re

árfolyamot.) Az opció ára ilyen esetben egy vegyes eloszlást követ. Értéke  $p$  valószínűséggel nulla, és  $(1 - p)$  valószínűséggel valamilyen folytonos eloszlás, amely hasonlatos a lognormális eloszláshoz.

### 6.1.3. Az opcióár becslése variancia csökkentő módszerek használatával

Tegyük fel, adott paraméter beállítás mellett az opcióár eloszlása  $\mu_X$  várható értékkel és  $\sigma_X$  szórással jellemzett eloszlás. Az opcióár egy realizációja, mely ténylegesen kialakul, nyilván nem azonos a  $\mu_X$ -vel, melyet a piaci szereplő felhasznál döntéséhez. Azt szoktuk mondani, ha végtelen sokszor kellene döntést hoznia, akkor biztosan a pénzénél lenne, így viszont bizonytalan az üzlet kimenetele. A bizonytalanság mértéke pedig  $\sigma_X$ -től függ. Emellett a döntéshozó valójában nem is  $\mu_X$  alapján dönt, hiszen azt csak becsülni

tudja, például az egyszerű mintaátlaggal ( $\bar{X}$ ). A döntés alapját tehát nem a valós paraméter képezi, és nem mindegy, hogy ez a ténylegestől milyen távol esik. Ez tehát  $\sigma_X^2$  mellett egy további bizonytalansági tényező, és pedig  $\text{var}(\bar{X}) = \sigma_X^2 / \text{SSIZE}$  mértékű. A becslés varianciája tehát a mintaméret növelésével csökkenthető, azonban mint említettük, a modell (többszöri) kiértékelése jelentős erőforrásigénnyel jár. Variancia csökkentő módszerek használatával azonban javíthatunk becslésünk tulajdonságain.

**73. Megjegyzés.** Amikor az opcióár eloszlásának paraméteréről beszélünk, akkor mindig a szórás fogalmát használjuk ( $\sigma_X$ ), megkülönböztetve a becslés varianciájától ( $\sigma_X^2 / \text{SSIZE}$ ).

Az ún. *kontroll változós* módszer egy sokváltozós eljárás, melynek alkalmazásához szükséges lesz a többdimenzós normalitás tesztek használatára. Ez képezi a jelen fejezetben bemutatott modellnek a disszertáció szempontjából releváns részét, mellyel demonstrálni fogjuk a normalitás hiányának egy lehetséges következményét. A becslőfüggvény ugyanis csak akkor marad torzítatlan, ha a módszer alapját képező többdimenziós normalitás szigorú hipotézise teljesül.

A torzítást empirikusan, Monte-Carlo szimuláció segítségével fogjuk meghatározni, mely azonban az igen nagy méretű replikációk ellenére kis mértékű mintavételi ingadozásnak van kitéve. Ahhoz, hogy a torzításról és a mintavételi hiba mértékéről fogalmunk legyen, bemutatunk egy másik variancia csökkentő módszert, és pedig a *korreláció indukciós* módszert, mely az alapeloszlástól függetlenül torzítatlan becslést eredményez. Az elméletileg zérus torzítást szintén Monte-Carlo szimulációval fogjuk kiszámítani, így láthatjuk majd, hogy milyen nagyságrendű véletlen hatással kell számolnunk a torzítás becslésénél.

**74. Megjegyzés.** Az opcióár varianciája továbbra is  $\sigma_X^2$  marad, mivel nem az opcióár eloszlását befolyásoljuk, hanem paramétereinek becsléséhez konstruálunk jobb becslő függvényt!

**A kontroll változós módszer.** Az időszak végi opcióár eloszlásának kvalitatív vizsgálata (12-13. ábrák) során azt feltételezzük, hogy közelítőleg lognormális eloszlást követnek, vagyis logaritmusuk követ normális eloszlást a  $K$  paraméter egy megfelelő tartományán. Az eloszlásnak, mint tudjuk, két paramétere van ( $\mu_x$  és  $\sigma_X$ ), esetünkben egyik sem ismert, mintából kell becsülnünk.

A 6.1. pont bevezető részében azt mondtuk, hogy minden szimulációs modellben nyilvánvalóan generálunk inputként véletlen változókat, melyek eloszlásának típusa és paraméterei ismertek. Emellett a modellen belüli transzformációk során keletkezh(et)nek további változók, melyek eloszlását illetve paramétereit meg tudjuk határozni. esetünkben pl. a kiinduló véletlen

számok a  $[0, 1]$  intervallum egyenletes eloszlásból származnak, ezekből transzformálunk standard normális változókat, melyből aztán előállítjuk a részvények idősorát, mely Wiener-folyamot követ. Mind az input, mind pedig a belőle származtatott változók valamilyen szinten korrelálnak az eredmény változóval, máskülönben a modell outputja független lenne az inputtól.<sup>17</sup> Ezt a függőséget kihasználhatjuk, ha ismerjük ezen változók (a továbbiakban: kontroll változók) eloszlását, annak típusát és paramétereit (ld. (12) összefüggést). Mi több, ha a kontroll változók és a modell eredmény változójának együttes eloszlása többdimenziós normális eloszlást követ, akkor lineáris modellt építhetünk.

A 1.1. pontban ismertettünk egy alapvető sokváltozós statisztikai összefüggést (11. tétel), mely azt állítja, hogy a többdimenziós normális eloszlás peremeloszlásai, valamint változóinak egymásra vonatkozó feltételes eloszlásai is normális eloszlásúak. Az egyszerűbb követhetőség miatt a tételt ehelyütt megismételjük.

Tekintsünk egy  $\mathbf{x} \sim N(\mu, \Sigma)$  többdimenziós normális eloszlást, és végezzük el a következő particionálásokat:

$$\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2], \quad \mu = [\mu_1, \mu_2], \quad \Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \quad (20)$$

**11. Tétel.** (Anderson [1958]): Ha  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2] \sim N(\mu, \Sigma)$ , akkor

$$\mathbf{x}_1 \sim N(\mu_1, \Sigma_{11}) \quad \text{valamint} \quad \mathbf{x}_2 \sim N(\mu_2, \Sigma_{22}),$$

$\mathbf{x}_1$ -nek  $\mathbf{x}_2$ -re vonatkozó feltételes eloszlása pedig szintén normális

$$\begin{aligned} \mu_{1|2} &= \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(\mathbf{x}_2 - \mu_2) \\ \Sigma_{11|2} &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} \end{aligned}$$

paraméterekkel.

**75. Megjegyzés.** Ha  $\Sigma_{12} \neq \mathbf{0}$ , akkor  $\Sigma_{11|2} \leq \Sigma_{11}$ , vagyis ha a két változócsoporthoz van korreláció, akkor  $\mathbf{x}_{1|2}$  változó varianciája kisebb, mint  $\mathbf{x}_1$  változóé.

Erre a tételre épül a következő lineáris variancia csökkentő módszer. Tételezzük fel, hogy a szimulációs modellünk egyik eredmény változójának ( $Y$ ) eloszlását kívánjuk jellemezni annak várható értékével, erre szeretnénk egy alacsony varianciájú, lehetőleg torzítatlan becslést konstruálni. Feltetésünk szerint  $Y$  közelítőleg normális eloszlást követ. A modellben kiválasztunk egy kontroll változót ( $C$ ) is, melyről viszont tudjuk, hogy normális eloszlást követ  $\mu_c$  és  $\sigma_c$  paraméterekkel.

<sup>17</sup>Itt most arra próbálunk kilyukadni, hogy ebben a modellben lineáris összefüggés van, ezért használjuk a függetlenség fogalmát. Zérus korreláció esetén valamilyen nemlineáris módszer után kellene néznünk.

**76. Megjegyzés.** A futamidő végi árfolyamok  $K$  bizonyos értékeire feltevésünk szerint lognormális eloszlást követnek, vagyis logaritmusukra tételezünk fel normális eloszlást. A 71. megjegyzés alapján viszont tudjuk, hogy a részvényárfolyamok  $t$  időpontbeli eloszlása szintén lognormális eloszlású, paramétereit pedig ismerjük, hiszen az árfolyamokat mi generáljuk.

Bevezetjük a következő becslőfüggvényt:

$$Y_{CV} = Y - \beta(C - \mu_C)$$

ahol  $Y$  a szimulációs modell (egyik) eredmény változója,  $C$  a kontroll változó, melynek ismert elméleti várható értéke  $\mu_C$ ,  $\beta$  pedig az ún. kontroll paraméter. Ez a becslés torzítatlan, hiszen

$$\mu_Y = E[Y_{CV}] = E[Y] - \beta E[(C - \mu_C)] = E[Y] \quad (21)$$

feltéve, hogy  $\beta$  konstans. A becslés akkor hatásos, ha  $Y_{CV}$  varianciája kisebb, mint  $Y$  varianciája. Belátható, hogy az *optimális kontroll* paraméter (Anderson [1958])

$$\beta = \frac{\sigma_{YC}}{\sigma_C^2} \quad (22)$$

ahol  $\sigma_{YC}$  az output változó és a kontroll változó kovarianciája,  $\sigma_C^2$  pedig a kontroll változó varianciája. A továbbiakban  $\sigma$ -val mindig az elméleti értéket,  $s$ -el pedig a mintából számított értéket jelöljük.

**77. Megjegyzés.** A (13) felírásnak esetünkben a (21) formula felel meg. Az elsőfajú regressziós becslést használjuk együttesen normális eloszlást követő változókra, melynek konkrét alakját a (11) tételből nyertük.

**78. Megjegyzés.** A gyakorlatban természetesen nem egyelemű mintából becsülünk. Az egyszerű mintaátlagra a (21) összefüggés értelemszerűen  $\bar{Y}_{CV} = \bar{Y} - \beta(\bar{C} - \mu_C)$ .

A variancia csökkenés mértéke annál nagyobb, minél nagyobb az eredmény változó és a kontroll változó közti korreláció, valamint minél kisebb a kontroll változó varianciája, ugyanis

$$Var(Y_{CV}) = \sigma_Y^2 + \beta^2 \sigma_C^2 - 2\beta \sigma_{YC}$$

A becslés akkor hatásos, ha  $\beta^2 \sigma_C^2 - 2\beta \sigma_{YC} \leq 0$ . Ekkor a (22) felhasználásával azt kapjuk, hogy

$$\frac{\sigma_{YC}}{\sigma_C^2} - 2\frac{\sigma_{YC}}{\sigma_C^2} \leq 0, \text{ azaz } \frac{\sigma_{YC}}{\sigma_C^2} = \rho_{YC}^2 \geq 0. \quad (23)$$



Mivel  $\rho_{YC}$  korrelációs együttható, ezért látható, hogy a variancia csökkentés feltétele a becslt és a kontroll változó közti korreláció.

Az  $Y$  és  $C$  együttes eloszlását nem ismerjük, így  $\sigma_{YC}$  sem ismert, ezért a kontroll paramétert mintából kell becsülnünk. Legyen

$$\hat{\beta} = b = \frac{s_{YC}}{\sigma_C^2}$$

Mivel a kontroll paramétert is mintából kell becsülnünk, ezért a szabadságfokok veszteséget indukálnak, a (23) feltétel megfelelője a

$$\rho_{YC}^2 \geq \frac{q}{n-2}$$

összefüggés, ahol  $q$  a kontroll változók száma,  $n$  pedig a mintaméret. A várható variancia csökkenés mértéke pedig

$$\eta = \frac{n-2}{n-q-2} (1 - \rho_{YC}^2) \quad (24)$$

**79. Megjegyzés.** Esetünkben  $q = 1$  és  $n = SSIZE$ .

Most viszont a (21) összefüggés már nem feltétlen igaz, hiszen  $b$  nem konstans, a minta függvénye. Tegyük fel, hogy  $(Y, C) \sim G_2(\mu_Y, \sigma_Y, \mu_C, \sigma_C, \sigma_{YC})$  ahol  $G_2$  kétdimenziós normális eloszlás. Ekkor igaz, hogy

$$\begin{aligned} \mu_Y &= E[Y_{CV}] = E[Y] - E[b(C - \mu_C)] = E[Y] - E\left[\frac{s_{YC}}{\sigma_C^2} (C - \mu_C)\right] \\ &= E[Y] - E\left[\frac{s_{YC}}{\sigma_C^2}\right] E[(C - \mu_C)] = E[Y] \end{aligned}$$

mivel a normális eloszlás átlagvektora és a minta kovariancia mátrixa független statisztikák (ld. például *Anderson[1958]*, *Móry-Székelly [1986]*).

Az előbbi összefüggést közvetlenül alkalmazhatnánk modellünkre, ha mind az eredmény változó (opcióár), mind pedig a kontroll változó ( $t$  időpontbeli részvényárfolyam) normális eloszlást követne.<sup>18</sup> Csakhogy a logaritmusuk követ normális eloszlást. A kontroll változó esetében ez nem okoz problémát, hiszen ha  $C' \sim LN(\mu_{C'}, \sigma_{C'})$  a  $t$ . időszaki árfolyam eloszlása, akkor  $C = \log(C') \sim N(\mu_C, \sigma_C)$  és  $\mu_C = \log(\mu_{C'}) - 1/2 \log(\sigma_{C'}^2/\mu_{C'}^2 + 1)$  valamint  $\sigma_C = \log(\sigma_{C'}^2/\mu_{C'}^2 + 1)$ , így a becsléshez fel tudjuk használni.

Az output változó esetében szintén elvégezhetjük az  $Y = \log(X)$  transzformációt, ahol  $X$  az opcióárat jelöli, és kiszámíthatjuk a  $Y_{CV}$  becslést. Az

<sup>18</sup>Ez persze csak az első lépcső. Nekünk ugyanis arra van szükségünk, hogy az együttes eloszlásuk legyen normális. Az együttes normalitásnak viszont a peremek normalitása szükséges feltétele.

$Y_{CV}$  egzakt eloszlását azonban már nem ismerjük<sup>19</sup>, így nem tudunk visszajutni az  $X$  változó terébe. Mi ui.  $\mu_X$  egy becslését akarjuk megadni, nem pedig  $\mu_Y$ -ét. Csakhogy

$$E[e^{Y_{CV}}] \neq \mu_X$$

legalábbis nem bizonyítható. Ezért megpróbálunk  $X$ -re vonatkozóan közvetlenül egy becslést előállítani.

Legyen  $Y = \log(X)$  továbbra is, valamint  $C = \log(C')$  és tekintsük a

$$X_{CV} = X - b(C - \mu_C) \quad (25)$$

becslést, ahol  $b = s_{YC}/\sigma_C^2$ ! Ez a korábban elmondottak alapján szintén torzítatlan becslés, mivel a kontroll paraméter ugyanaz, vagyis  $E[X_{CV}] = \mu_X$ . Az viszont nem biztos, hogy  $b$  illetve  $\beta$  most is optimális, ugyanis

$$\text{Var}(X_{CV}) = \sigma_X^2 + \beta^2 \sigma_C^2 - 2\beta \sigma_{XC} \quad \text{ahol } \beta = \frac{\sigma_{YC}}{\sigma_C^2}$$

variancia csökkenés pedig akkor várható, ha

$$\begin{aligned} \frac{\sigma_{YC}^2}{\sigma_C^2} - 2 \frac{\sigma_{YC}}{\sigma_C^2} \sigma_{XC} &\leq 0 \\ \frac{\sigma_{YC}^2}{\sigma_C^2} - 2 \frac{\sigma_{YC}}{\sigma_C^2} \frac{\sigma_{YC}^2}{\sigma_{YC}^2} \sigma_{XC} &= \frac{\sigma_{YC}^2}{\sigma_C^2} \left(1 - 2 \frac{\sigma_{XC}}{\sigma_{YC}}\right) \leq 0 \\ \frac{1}{2} &\leq \frac{\sigma_{XC}}{\sigma_{YC}} \end{aligned}$$

Ezek szerint variancia csökkenés abban az esetben várható, ha  $X$  és  $C$  kovarianciája legalább fele akkora, mint  $Y$  és  $C$  kovarianciája. Belátható, hogy a variancia csökkentés akkor maximális, ha a fenti hányados éppen egységnyi, azaz  $\sigma_{YC} = \sigma_{XC}$ , tehát  $\beta = \sigma_{XC}/\sigma_C^2$  az optimális kontroll.

Mivel  $\sigma_{XC}$  sem ismert, ezért annak becslését kell használnunk, jelöljük ezt  $s_{XC}$ -vel. A torzítatlansághoz azt kellene belátnunk, hogy  $s_{XC}$  és  $C$  (illetve  $\bar{C}$ ) függetlenek. Ez viszont igaz, hiszen a  $s_{XC} = f(s_{YC})$  leképezés létezik, bijektív és folytonos is. Gondoljuk meg ugyanis, hogy az empirikus kovariancia ( $s_{YC}$ ) valójában egy leképezés az  $Y^n \times C^n$  mintatérből a valós számegyenesre (ahol  $n$  a mintaméret). Az  $s_{XC}$  leképezés esetében pedig az  $n$  darab  $Y$  tengely helyett azok szigorúan monoton transzformációját,  $X = e^Y$  tengelyeket használjuk. Tehát a két leképezés szintfelületei bijektív viszonyban állnak egymással. Ekkor pedig a valószínűségi változók függvényeire vonatkozó tétel alapján mivel  $s_{YC}$  és  $C$  is független, ezért  $s_{XC}$  és  $C$  is független, így a

$$b = \frac{s_{XC}}{\sigma_C^2} \quad (26)$$

kontroll nemcsak optimális, de torzítatlan becslést is eredményez.

<sup>19</sup>Ha a kontroll paraméter ( $\beta$ ) konstans lenne, akkor a (11) tétel alapján  $Y_{CV}$  eloszlása normális lenne. A helyzet viszont az, hogy a kontroll paramétert mintából kell becsülni.

**80. Megjegyzés.** Ez a gondolatmenet nyilván nem csak a lognormális eloszlás esetében igaz. Minden olyan eloszlás esetében használható, mely normális eloszlásból származtatható. Ezért a normalitást, illetve többdimenziós normalitást megkövetelő módszerek a normális eloszlásra, illetve adott esetben a lineárisra visszavezethető modellek esetében is használhatóak maradnak.

A (25) lineáris becslést fogjuk felhasználni, ahol  $b$  a (26) szerint határozódik meg. A modell eredmény változóját, vagyis az opcióárat  $X$ -szel jelöltük, és feltevésünk szerint lognormális eloszlást követ ismeretlen  $\mu_X$  és  $\sigma_X$  paraméterekkel. Definiáltuk továbbá  $Y$ -t, mint  $X$  logaritmusát, mely, ha a feltevésünk helyes, normális eloszlású. Kontroll változónak választhatjuk valamely  $t$  (például  $t = T$ ) időpontra vonatkozó részvényárfolyamot ( $C'$ ), pontosabban annak logaritmusát ( $C = \log(C')$ ), melyről viszont tudjuk, hogy normális eloszlású  $\mu_C$  és  $\sigma_C$  paraméterekkel. Azt kell most már csak belátnunk, hogy a két változó együttes eloszlása is normális, vagyis létezik a (20) struktúra, hiszen a kontroll paramétert mintából kell becsülnünk. Összefoglalva tehát a következő helyzet állt elő:

$$\begin{aligned} Y &\sim \hat{N}(\hat{\mu}_Y, \hat{\sigma}_Y) \\ C &\sim N(\mu_C, \sigma_C) \\ (Y, C) &\sim \hat{N}_2(\hat{\mu}_Y, \hat{\sigma}_Y, \mu_C, \sigma_C, \hat{\sigma}_{YC}) \end{aligned}$$

ahol kalapos betűk jelzik, hogy a kérdéses érték illetve eloszlás ismeretlen, a változó értékét becsüljük, illetve az eloszlásra vonatkozóan feltevéssel élünk.

**81. Megjegyzés.** Ennél a módszernél tehát kihasználjuk a normális eloszlás által implikált linearitást és a minta statisztikák függetlenségét. Ugyanakkor ha  $(Y, C)$  együttes eloszlás normális, akkor definíció szerint  $(X, C')$  kétdimenziós lognormális eloszlás, tehát a két eloszlástípus használatához gyakorlatilag azonos az eszköztárunk.

Célunk elsősorban az, hogy valós adatokon is összehasonlítsuk az általunk vizsgált tesztek, valamint hogy érzékeltesük a normalitás hiányának következményeit a gyakorlatban. Mivel a 5.1. részben bemutatott  $F$ ,  $L$ ,  $N$  és  $T$  tesztek hasonló eredményeket adtak, ezért az összehasonlításhoz csak az  $F$  variánst választottuk ki. A vizsgálatba természetesen bevontuk a  $W$  és a  $H$ -tesztet.

A kontroll változós módszer esetében a normalitás hiányának valószínű következménye az lesz, hogy torzított becsléseket kapunk.<sup>20</sup> A torzítások mértékét empirikusan, Monte-Carlo szimulációval becsülhetjük a következőképpen. Futtassuk le a modellt valamely paraméterbeállítás mellett, és

<sup>20</sup>A normalitás, mint említettük, ahhoz is szükséges, hogy lineáris becslést használhassunk. Ezzel a kérdéssel ehelyütt nem kívánunk foglalkozni.

legyen a mintaméret  $SSIZE$ . Számítsuk ki az opcióárak egyszerű átlagát ( $\bar{X}^{(1)}$ ), valamint a (21) által megadott korrigált becslést ( $\bar{X}_{CV}^{(1)}$ ). Ismételjük meg a kísérletet  $r$ -szer különböző, egymástól függetlenül vett mintákra, így a  $\bar{X}^{(1)}, \dots, \bar{X}^{(r)}$  illetve  $\bar{X}_{CV}^{(1)}, \dots, \bar{X}_{CV}^{(r)}$  becsléseket kapjuk. Az elméleti torzítás  $\hat{\mu}_X - \mu_X$ , vagyis a becsült paraméter várható értékének és a valós paraméternek a különbsége. Empirikusan a várható értéket ebben az esetben is a mintaátlaggal becsüljük, ahol a mintaméret most  $r$ , a replikációk vagy makro-replikációk száma. Míg  $SSIZE$  tipikusan egy alacsony érték, addig  $r$ -et magasnak választjuk meg, hiszen elméleti paramétereket akarunk empirikusan vizsgálni. A  $\mu_X$  torzítatlan becslése az egyszerű mintaátlagok átlaga, ami pedig az összes  $r \cdot SSIZE$  darab becslés egyszerű átlagával azonos. Itt tehát az effektív mintaméret igen magas. Ennek megfelelően 
$$\hat{\mu}_X = 1/r \sum_{k=1}^r \bar{X}_{CV}^{(k)}.$$

A 7. táblázat  $F$ ,  $W$  és  $H$  oszlopa közli, hogy egy adott  $K$  paraméter mellett a többdimenziós Shapiro-Wilk tesztek elfogadják-e a kontroll változó és a modell eredmény változójának együttes normalitására vonatkozó feltevést. Minden egyes replikáció alkalmával elvégezve az egyes teszteket, számláljuk az elutasítások számát  $\alpha = 0.10, 0.05, 0.01$  szignifikancia szinteken. A normalitás akkor fogadható el, ha az elutasítási ráták rendre megegyeznek a megadott szignifikancia szintekkel, vagyis például  $\alpha = 0.05$  szinten a mintáknak nagyjából az 5%-át utasítjuk el, hiszen ennyi az elsőfajú hiba mértéke. Ha az elutasítások mértéke ennél nagyobb, akkor a minták nem normális eloszlásból jönnek.

7. táblázat. Kontroll változós módszer

$K$	$F$	$W$	$H$	$\sigma^2$	$\sigma_{CV}^2$	$\eta$	$Torzítás$	$MSE$	$MSE_{CV}$	$\eta_{MSE}$	$\delta$	$\eta_0$
0.0	0.2522 <b>0.1558</b> <i>0.0467</i> 0.0436	0.2686 <b>0.1662</b> <i>0.0500</i> 0.2578	0.1001 <b>0.0486</b> <i>0.0092</i> 0.0996	0.0201	0.0177	88.04	0.0068	0.0201	0.0178	88.27	0.40	85.79
1.0	<b>0.1488</b> <i>0.0444</i> 0.2255	<b>0.1579</b> <i>0.0464</i> 0.2296	<b>0.0474</b> <i>0.0091</i> 0.0972	0.0202	0.0178	88.07	0.0066	0.0202	0.0179	88.39	0.40	85.79
2.0	<b>0.1350</b> <i>0.0395</i> 0.1736	<b>0.1405</b> <i>0.0405</i> 0.1796	<b>0.0470</b> <i>0.0101</i> 0.0985	0.0205	0.0181	88.31	0.0062	0.0205	0.0182	88.50	0.39	86.59
4.0	<b>0.0993</b> <i>0.0272</i> 0.1213	<b>0.1028</b> <i>0.0259</i> 0.1286	<b>0.0480</b> <i>0.0091</i> 0.1050	0.0207	0.0184	88.53	0.0051	0.0207	0.0184	88.66	0.39	86.59
8.0	<b>0.0657</b> <i>0.0168</i> 0.1305	<b>0.0722</b> <i>0.0176</i> 0.1361	<b>0.0573</b> <i>0.0127</i> 0.1161	0.0212	0.0192	90.45	0.0017	0.0212	0.0192	90.46	0.38	87.38
10.0	<b>0.0737</b> <i>0.0211</i> 0.2758	<b>0.0773</b> <i>0.0202</i> 0.3252	<b>0.0622</b> <i>0.0151</i> 0.1847	0.0210	0.0200	91.33	-0.0008	0.0210	0.0201	91.34	0.38	87.38
16.0	<b>0.1876</b> <i>0.0702</i> 0.6985	<b>0.2199</b> <i>0.0819</i> 0.8053	<b>0.1137</b> <i>0.0404</i> 0.5977	0.0278	0.0260	93.84	-0.0073	0.0278	0.0261	94.04	0.33	91.01
32.0	<b>0.6068</b> <i>0.4089</i> 0.7968	<b>0.7063</b> <i>0.4644</i> 0.9850	<b>0.5101</b> <i>0.3312</i> 0.7085	0.0426	0.0422	98.98	-0.0209	0.0426	0.0426	100.00	0.25	95.74
50.0	<b>0.6858</b> <i>0.4449</i>	<b>0.9685</b> <i>0.8943</i>	<b>0.5449</b> <i>0.2204</i>	0.1292	0.1401	108.34	-0.0554	0.1292	0.1431	110.74	0.11	100.89

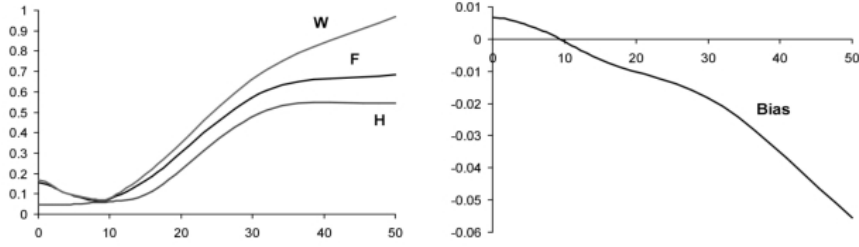
Az egyes szignifikancia szintekhez tartozó elutasítási hányadosokat az áttekinthetőség kedvéért különböző típusú betűkkel szedtük. A  $\alpha = 0.10$  szintet normál,  $\alpha = 0.05$  szintet félkövér,  $\alpha = 0.01$  szintet pedig dőlt betűvel szedtük. A táblázatban közöljük ezenfelül az egyszerű mintaátlag ( $\sigma^2$ ) illetve a korrigált mintaátlag varianciáit ( $\sigma_{CV}^2$ ), valamint a torzítást és a legkisebb négyzetes hibákat ( $MSE$ ,  $MSE_{CV}$ ). Ahol a tesztek alapján úgy döntünk, hogy elfogadjuk a normális eloszlás hipotézisét, vagyis torzítatlan a becslésünk, ott a variancia csökkentés mértéke a modell értékelésének kritériuma, egyéb esetben a legkisebb négyzetes hibában elért csökkenés. A táblázat  $\rho$  oszlopa tartalmazza a becsült korrelációs együtthatókat,  $\eta_0$ ,  $\eta$  és  $\eta_{MSE}$  oszlopok pedig a várható variancia/legkisebb négyzetes hibában bekövetkező csökkenés mértékét:

$$\begin{aligned}\eta_0 &= (SSIZE - 2) / (SSIZE - 3) (1 - \rho^2) \quad (\text{elméleti}) \\ \eta &= \sigma_{CV}^2 / \sigma^2 \\ \eta_{MSE} &= MSE_{CV} / MSE\end{aligned}$$

A táblázatból három fontos következtetés vonható le. Először is a  $W$ -teszt mindvégig szigorúbbnak bizonyult, mint a másik kettő. Sikertelt tehát olyan alkalmazást találunk, ahol ezen teszt előnyei hasznosíthatók.

Másodszor mind az  $F$ , mind pedig a  $W$ -teszt nagyjából a  $K \in [8.0, 10.0]$  tartományban fogadja el leginkább a normalitást. Az  $\alpha = 0.05$  szinten például (félkövér betű)  $F$ -teszt értékei  $0.0657 - 0.0737$ , a  $W$ -teszt értékei pedig  $0.0722 - 0.0773$ . Ezen a tartományon kívül mindenhol magasabbak az elutasítási ráták. Ha az empirikus torzításokra pillantunk, akkor láthatjuk, hogy ebben a tartományban a legalacsonyabbak ( $0.0017$  és  $-0.0008$ ). A  $K \in [0.0, 8.0]$  esetén a becslés felfelé, a  $K \in [10.0, 50.0]$  tartományban pedig lefelé torzít.

A harmadik fontos észrevétel, hogy a  $H$ -teszt esetében az elutasítási ráták nagyjából 5% körül mozognak ( $0.47 - 0.62$ ) a  $K \in [0, 10.0]$  tartományban, majd efelett  $K$  növekedésével monoton növekednek. Vagyis csak a  $K \in [10.0, 50.0]$  tartományban utasítja el az együttes normalitás hipotézisét, és itt is gyengébbnek bizonyul a többinél. Ez abból fakad, hogy az opcióár logaritmus kis  $K$  értékekre közelítőleg normális eloszlású, majd  $K$  növekedésével egyre inkább elfajul (ld. 12-13 ábrák). A  $H$ -teszt erősen a peremekre épít, így ezt az elmozdulást képes csak érzékelni, az együttes eloszlás struktúrájában bekövetkező változást nem. A példa jól illusztrálja, hogy habár a  $K = 0$  esetén a legszebbek a peremek, az együttes normalitáshoz hasonlatos struktúrát mégis csak a  $K \in [8.0, 10.0]$  értékek esetén találunk. A tesztek eredményeit a kontroll változós módszeren végzett torzításvizsgálat is alátámasztja, vagyis a módszer tényleg ott működik a legjobban, ahol a teszt az alapeloszlást a leginkább normálisnak értékeli. Megjegyezzük, hogy ahol az együttes normalitás teljesül, ott ezt a  $H$ -teszt is elfogadja a peremekre alapozva, tehát az eredmények konzisztensek.



14. ábra. Többdimenziós normalitás tesztek és empirikus torzítás

A 14. ábrán az illusztráció kedvéért grafikonon is szemléltetjük eredményeinket. A bal oldali grafikon mutatja a próbafüggvények empirikus értékeit a három teszt esetében. Látható, hogy az  $F$  és  $W$ -tesztek U-alakú görbék, míg a  $H$ -teszt fordított L-alakú. Az is megfigyelhető, hogy a  $W$ -teszt jelentősen dominálja a másik kettőt. A 7. táblázatból kiolvasható, hogy pl.  $K = 50$  esetén a  $W$ -teszt már majdnem biztosan elutasítja a normalitást (0.9685), az  $F$ -teszt csak az esetek 70%-ában (0.6858), a  $H$ -teszt pedig alig több, mint az esetek felében (0.5449). A jobb oldali grafikon a torzítás alakulását mutatja teljes összhangban a tesztek eredményeivel. A kontroll változós módszer nyilvánvalóan érzékeny az együttes normalitás hiányára, a normalitás pedig a  $K$  értékétől függ érzékenyen.

Az empirikus torzítások értékelésénél számításba kell venni a mintavételi hibát, illetve a véletlenszám-generátor hibáját is, ami harmadik és negyedik tizedesjegyen már jelentkezik. Ahhoz, hogy ennek mértékéről fogalmunk legyen, megvizsgáltuk a modellt egy másik variancia csökkentő módszerrel.

**A korreláció indukciós módszer.** A független azonos eloszlású mintát az egyszerű kezelhetősége miatt használjuk, azonban korrelált mintából általában hatásosabb becslést készíthetünk. Ha a várható értéket kívánjuk becsülni az egyszerű mintaátlaggal, akkor nem a FAE minta a legjobb választás, mivel annak a legkisebb az információtartalma. A várható értékre vonatkozó  $E[\sum X_i] = \sum E[X_i]$  tétel ugyanis általánosan igaz, nem szükséges a mintaelemeket egymástól függetlenül mintavételezni, ezért  $E[\bar{X}] = \mu_X$  is igaz korrelált mintákra.

Az opcióár modellt újra futtatuk, de a modell bemenő adatait nem egymástól függetlenül generáltuk, hanem az ún. *korreláció indukciós* módszerrel mintavételeztük. Az előzőekben úgy generáltuk a bemenő adatokat, hogy a  $[0, 1]$  intervallumon egyenletes eloszlású, független  $U_1^{(1)}, \dots, U_T^{(1)}$  kiinduló adatokat generáltunk, ahol  $t = 1, \dots, T$  az egyes időszakokat jelenti. Ez volt a modell inputja, ezeket normális eloszlású változókká transzformáltuk, majd a (19) formula alapján létrehoztuk a részvényárfolyamok egy szcenárióját az adott hosszúságú futamidőre, majd a modellt futtatva

kiszámítottuk az opcióárat. Ezt *SSIZE*-szor ismételtük meg, mégpedig oly módon, hogy az  $i+1$ -dik lépésben az  $U_1^{(i+1)}, \dots, U_T^{(i+1)}$  input adatokat az előző  $U_1^{(i)}, \dots, U_T^{(i)}$  adatoktól függetlenül generáltuk. Vagyis az egyes szcenáriókból kiszámított opcióárakból egy *SSIZE* méretű FAE mintát nyertünk, és ebből a mintából becsültük az opcióár várható értékét. A korreláció indukciós módszer (*CI*) esetében az egyes szcenáriókat nem egymástól függetlenül generáljuk, amint azt a (27) séma mutatja.

$$\left. \begin{array}{ccc} \overbrace{U_1^{(1)}, \dots, U_T^{(1)}}^{\text{független változók}} & \rightarrow & X_1 \\ & & \vdots \\ & & \vdots \\ U_1^{(SSIZE)} \dots U_T^{(SSIZE)} & \rightarrow & X_{SSIZE} \end{array} \right\} \text{korrelált minta} \quad (27)$$

A korreláció indukció megértéséhez mindenekeelőtt megismerkedünk a negatív síknegyedben való függőség (*n.q.d.*) fogalmával.

**82. Definíció.** Az  $(A_1, A_2)$  kétdimenziós véletlen változó vektor akkor *n.q.d.*, ha

$$\Pr(A_1 \leq a_1, A_2 \leq a_2) \leq \Pr(A_1 \leq a_1) \Pr(A_2 \leq a_2)$$

*tetszőleges  $a_1$  és  $a_2$ -re*

Ha egyenlőség áll fenn, akkor függetlenek. Az *n.q.d.* koncepciójára alapozva bevezetjük az eloszlások egy  $\mathcal{G}$  osztályát, melyet a szimulációs modellek véletlen változó inputjaként fogunk felhasználni.

**83. Definíció.** Legyen  $G^{(SSIZE)} \in \mathcal{G}$ , ekkor

$CI_1 : SSIZE \geq 2$ -re  $G^{(SSIZE)}$  egy *SSIZE* dimenziós eloszlás  $U(0, 1)$  eloszlású peremekkel

$CI_2 : G^{(SSIZE)}$  mindegyik kétdimenziós peremeloszlása *n.q.d.*

A fenti definíció alapján megadjuk a (27) szerinti mintavételi tervet. Legyen tehát

$$X_i = \varphi(U_1^{(i)}, \dots, U_T^{(i)}) \quad i = 1, \dots, SSIZE \quad (28)$$

a  $i$ -dik szcenárió outputja, ahol  $\varphi(\cdot)$  a szimulációs modell által megvalósított leképezés, valamint legyen

$$\mathbf{U}_t = [U_t^{(1)}, \dots, U_t^{(SSIZE)}] \quad t = 1, \dots, T \quad (29)$$

vagyis a (27) sémában a  $t$ -dik oszlop. A mintavételi terv ekkor

- $SC_1$  : az  $\mathbf{U}_t$  véletlen vektor eloszlása  $G^{(SSIZE)} \in \mathcal{G}$ ,  $t = 1, \dots, T$
- $SC_2$  : az  $\mathbf{U}_1, \dots, \mathbf{U}_T$  vektorok páronként függetlenek



Az  $SC_1$  feltétel szerint a  $U_t^{(1)}, \dots, U_t^{(SSIZE)}$  változók (vagyis a sorok) páronként a negatív síknegyedben összefüggőek ( $CI_1$  tulajdonság), az  $SC_2$  pedig a  $CI_2$  tulajdonság miatt biztosítja, hogy a  $U_1^{(i)}, \dots, U_T^{(i)}$  változók (oszlopok) a  $[0, 1]$  intervallumon egyenletes, független azonos eloszlású változók legyenek.

**84. Definíció.** Az  $\{X_i, i = 1, \dots, SSIZE\}$  mintát  $G^{(SSIZE)}$ -mintának nevezzük, ha a (28) és (29) felírások alapján az  $SC_1$  és  $SC_2$  feltételek által meghatározott mintavételi terv szerint generáltuk.

A következő két állítás arra vonatkozik, hogy a szimulációs modell eredmény változójában valóban variancia csökkenés következik be.

**85. Következmény.** Ha  $G^{(SSIZE)} \in \mathcal{G}$ , továbbá  $\{X_i : i = 1, \dots, SSIZE\}$   $G^{(SSIZE)}$ -minta és  $\varphi(\cdot)$  minden argumentumában monoton függvény, akkor  $[X_i, X_j]_{i \neq j}$  n.q.d..

**86. Következmény.** Ha a  $(A_1, A_2)$  kétdimenziós vektor n.q.d., akkor

$$Cov(A_1, A_2) \leq 0,$$

és egyenlőség akkor és csak akkor áll fenn, ha  $A_1$  és  $A_2$  függetlenek.

A 85. következmény valójában az 1. tétel, a 86. következmény pedig a 3. lemma következménye Lehmann cikkében (Lehmann [1966]).

Mivel

$$Var \left[ \sum_{i=1}^{SSIZE} X_i \right] = \sum_{i=1}^{SSIZE} Var[X_i] + 2 \sum_{i \neq j} Cov[X_i, X_j] \leq \sum_{i=1}^{SSIZE} Var[X_i],$$

ezért az átlag becslés varianciája valóban kisebb lesz, mint FAE minta esetén, mégpedig a páronkénti kovarianciák összegével.

A  $\varphi(\cdot)$  leképezés monotonitásához biztosításához módosítani kellett a normális eloszlású véletlen szám generálást. A *gasdev* algoritmus, mint ismertettük, elvetésem módszert használ, vagyis adott bemenő  $U_1^{(i)}, \dots, U_T^{(i)}$  változók és a keletkező normális eloszlású  $Z_1^{(i)}, \dots, Z_T^{(i)}$  változók, s így a  $X_i$  eredmény változó között nincs függvnyszerű kapcsolat. Ehelyett az egyszerűbb  $Z = \Phi^{-1}(U)$  transzformációt használtuk, ahol az  $U$  egyenletes eloszlásból származó véletlen számból a  $\Phi^{-1}(\cdot)$  függvénnyel, a normális eloszlás inverz eloszlásfüggvényével képeztünk normális eloszlású változót. Elméletben ugyan ez lenne a triviális módszer, de mint tudjuk,  $\Phi^{-1}(\cdot)$  zárt alakban nem ismert, csak polinom közelítését tudjuk használni. Modellünkben az AS241 algoritmust használtuk fel (Wichura [1988]).

A korreláció indukciós módszernél gyakori és kedvelt mintavételi technika a *latin hiperkocka* (LHS) mintavételezés, melyet mi is használtunk. Generáljunk a  $[0, 1]$  intervallumon független egyenletes eloszlású  $U_t^{(i)}$  ( $t =$

$1, \dots, T$  és  $i = 1, \dots, SSIZE$ ) véletlen számokat, majd pedig képezzük az alábbi transzformációt:

$$\begin{aligned}\tilde{U}_t^{(i)} &= \frac{\pi_t(i) - 1 + U_t^{(i)}}{SSIZE} \\ U_t^{(i)} &\sim U(0, 1) \text{ i.i.d., } t = 1, \dots, T \text{ és } i = 1, \dots, SSIZE\end{aligned}\tag{30}$$

ahol  $\pi_t(1), \dots, \pi_t(SSIZE)$  az első  $SSIZE$  egész szám egy véletlen permutációja. Vesszünk tehát  $T$  darab egymástól független permutációt, és elvégezve az előbbi transzformációt a kívánt tulajdonságú  $\tilde{U}_t^{(i)}$  ( $t = 1, \dots, T$  és  $i = 1, \dots, SSIZE$ ) input adatokat nyerjük. Az *LHS* technika hasonlít az *EV* (egyszerű véletlen minta) mintavételhez abban, hogy a  $\pi_t(i)$  változókat a  $S = \{1, \dots, SSIZE\}$  halmazból visszatevés nélkül véletlenszerűen vesszük. Emellett viszont a  $\pi_1(i), \dots, \pi_T(i)$  értékeket egymástól függetlenül visszatevéssel mintavételezzük.

**87. Állítás.** A (30) formula alapján generált  $G_{LHS}^{(SSIZE)}$  minta a  $\mathcal{G}$  családhoz tartozik.

A módszert behatóbban vizsgálta McKay et al.[1979], Stein [1987] és Owen [1992a, b].

Az opcióár modellt tehát újrafuttattuk, de a kiinduló egyenletes eloszlású véletlen számokat latin hiperkocka mintavételezéssel generáltuk. A vizsgálat eredményét a 8. táblázatban közöljük <sup>21</sup>. Az opcióár eloszlásától függetlenül tehát elméletileg torzítatlan becslést kapunk. Célunk most azt vizsgálni, hogy az elméletileg zérus torzításokat választott  $r = 1000$  makro-replikációk mellett<sup>22</sup> mekkora hibával közelítjük. Ez alapján szét tudjuk választani a kontroll változós módszerrel kapott eredményeinkben a mintavételi hibát a módszer által implicált torzítástól.

Az empirikus torzításokra pillantva itt nem látható az a szisztematikus változás, ami a kontroll változós becslési módszer esetén tapasztalható: először csökkenés nagyjából  $K = 10$ -ig, majd pedig — abszolút értékben — ismét növekszik a torzítás. A becsült torzítások előjele a korreláció indukciós módszernél teljesen véletlenszerű, míg a kontroll változós becslés esetén alacsony  $K$  értékekre felfelé, magasabb  $K$  értékekre pedig lefelé torzít a becslés. A hibák nagyságrendje is jóval alacsonyabb, a legtöbb érték csak a negyedik tizedes jegyén szignifikáns.

A módszer kétségtelen előnye, hogy eloszlásfüggetlen, vagyis nincs szükség a többdimenziós normalitás alaphipotézisére ahhoz, hogy torzítatlan becslést kapjunk. Az is látható, hogy a Monte-Carlo szimuláció eredményei alapján

<sup>21</sup>Jelmagyarázat a másik két táblázatával azonos, kivéve, hogy itt  $\sigma_{CI}$  oszlop szerepel, a korreláció indukciós módszerrel becsült átlag varianciája. Tekintve, hogy ez a módszer elméletileg nem torzított becslést eredményez, az *MSE* oszlopok hiányoznak. Nem szerepel továbbá — értelemszerűen — a korrelációra vonatkozó becslés sem.

<sup>22</sup>Vagyis 1000 db becslést számolunk ki.

8. táblázat. Korreláció indukciós módszer

$K$	$\sigma^2$	$\sigma_{CI}^2$	$\eta$	Torzítás
0.0	0.0202	0.0102	50.61	-0.000514
2.0	0.0207	0.0105	50.86	0.000250
4.0	0.0211	0.0107	50.58	-0.002393
8.0	0.0226	0.0120	53.01	0.000137
10.0	0.0236	0.0129	54.74	-0.000745
16.0	0.0260	0.0183	70.45	-0.000012
32.0	0.0441	0.0353	80.07	-0.007234
50.0	0.1356	0.1121	90.05	0.006427

igen jelentős mértékű variancia csökkentést (50%!) tudtunk elérni. A problémát mégis az okozza, hogy a variancia csökkenés mértékét analitikusan (ex-ante) nem lehet meghatározni. Ha pl.  $SSIZE = 1000$  méretű mintából megbízhatóan tudnánk a várható értéket becsülni, de ez kínosan hosszú futási időt eredményez, akkor a kontroll változós módszer esetén a (24) formula alapján meghatározhatjuk, hogy mondjuk  $SSIZE = 800$  ugyanezt a pontosságot biztosítja. A korreláció indukciós módszer esetén csak annyit tudunk (éspedig a 86. következmény alapján), hogy a variancia biztosan alacsonyabb lesz adott  $SSIZE = 1000$  méretű mintából számított becslésre, de nem tudunk spórolni a futási idővel.

További probléma, hogy az opcióár eloszlásának a másik paramétere, a szórás ( $\sigma_X$ ) is érdekel bennünket. A korreláció indukciós módszernél azonban módosítjuk az input adatokat, így az opcióárra olyan mintát kapunk, melyből a várható érték még torzítatlanul becsülhető, a szórás azonban már nem. A kontroll változós módszer használata esetén  $\sigma_X$  továbbra is torzítatlanul becsülhető a korrigált tapasztalati szórással.

Megjegyezzük, hogy a kontroll változós módszerre vonatkozóan léteznek aszimptotikus tételek, vagyis melyek nagy minták esetén jól működnek (Nelson [1990]). A nagy alatt egy olyan mintaméretet kell érteni, mely az aszimptotikus tulajdonságokhoz elégséges, de még elfogadható ráfordítás mellett kiszámolható modellt eredményez. A normális eloszlás hipotézisére épülő modellek világát tehát a kutatók igyekeznek tágítani, ameddig csak lehetséges. Sajnos azonban vannak olyan területek, ahol a piac teljesen más mögöttes logika szerint működik, így a modellek építéséhez új módszertanra van szükség. Ilyen terület a piaci kockázatkezelés, melyre példát a most következő részben mutatunk be.

## 6.2. Kockázatotott érték számítás

A modern piaci kockázاتمérés legnépszerűbb elemzési rendszerét a *kockázatotott érték* — Value at Risk; *VaR*— számításához kapcsolódó módszerek jelentik. Ez tipikusan egy olyan alkalmazási terület, ahol az alapeloszlásra —

esetünkben a piaci hozamok eloszlására — vonatkozó normalitás hipotézise legkevésbé állja meg a helyét. Az alapadatok komponensenkénti, egydimenziós idősoraira többnyire a csúcsosság, illetve a hosszan elnyúló, vastag szélek a jellemzők.

A piaci hozamok idősorát lehet statikus illetve dinamikus modellel leírni. A statikus esetben feltételezzük, hogy az adatok varianciája minden időpontban azonos, míg a dinamikus esetben a variancia valamilyen sztochasztikus folyamat (pl. ARCH, GARCH) szerint változik. Mi most a statikus esettel foglalkozunk, vagyis igyekeztünk olyan idősort találni, ami a varianciára vonatkozó feltevésnek megfelel. Az itt közölt elemzés egy előkészületben lévő tanulmány része, melynek magam is társszerzője vagyok (Benedek *et.al.*[2001]).

A portfólió diverzifikációja a kockázatkezelésben kiemelkedő fontossággal bír, erre Markovitz tanulmánya hívta fel a figyelmet (Markovitz [1959]). A diverzifikáció úgy működik, hogy ha a portfóliót több kockázati faktor között allokaljuk, akkor habár egyes faktorok kedvezőtlen elmozdulása miatt veszíthetünk, más faktorok gyenge vagy negatív korreláltságuk miatt kompenzálják az előbbi veszteséget. Az, hogy meghatározott valószínűségi szinten a normalitás körülményeihez képest jóval nagyobb veszteség keletkezhet, mára már szervesen beépült a kockázatkezelők tudatába. Azonban azt is világosan kell látni, hogy a tőkepiaci válságok esetén, ami egy egész szektort vagy földrajzi régiót érint, az egyes faktorok együttesen produkálhatnak hatalmas veszteséget. Másszóval portfóliók esetén a legsúlyosabb problémát nem az okozza, hogy egy faktor extrém viselkedést produkál, hanem az, hogy az extremitást együttesen, egyszerre képesek mutatni, és ekkor a diverzifikáció már nem jelent akkora védelmet, mint azt normális piaci körülmények közt várnánk. Persze nem kell feltétlenül ritkán bekövetkező krach-okra gondolni. A mindennapok során az együttesen bekövetkező veszteségek gyakorisága jóval nagyobb, mint azt az elliptikus kontúrok alapján gondolhatnánk.

Ahhoz, hogy a részvényhozamok együttmozgását jobban megértsük, és a kockázatot minél valószerűbben meg tudjuk ragadni, már nem elégséges a lineáris korreláció fogalma. Ennek legfőbb oka az, hogy a lineáris korreláció csak akkor alkalmazható, ha az alapeloszlás elliptikus. Ezen belül pedig csak a többdimenziós együttes normális eloszlás esetében elégséges a kapcsolatszorosság mérésére, hiszen csak ebben az esetben írja le teljeskörűen az eloszlás függőségi viszonyait (ld. 29. illetve 66. megjegyzések).

Ha az alapeloszlás nem teljesíti az együttes normalitás követelményét, akkor alternatív függőségi mérőszámokra van szükség. Ilyen lehet például az *aszimptotikus függőség* (37. definíció), mely a tőzsdei extrém események mérésére alkalmasnak látszik. Lehetséges ugyanis, hogy a napi mozgásokat tekintve az egyes papírok között közepes vagy gyenge korrelációt mérünk, mégis, ha valami rendkívüli esemény történik, ezek a papírok egyszerre kezdenek zuhanni vagy emelkedni. Vagyis extrém esemény (*tail event*) bekövetkezése az egyik változóban együttjár egy másik változó értékében bekövet-

kező extremitással. Ebből a szempontból a normális eloszlás elég kedvezőtlen modellválasztás lenne, hiszen tudjuk a 38. megjegyzés alapján, hogy az eloszlás aszimptotikusan független. Ugyanakkor a Student  $t$ -eloszlás például, különösen alacsony szabadságfokoknál, aszimptotikusan összefüggő. A 9. táblázatot az *Embrechts et al. [1999]* cikkből vettük át.

9. táblázat. A  $t$ -kopula aszimptotikus függősége

$\nu/r_{ij}$	-0.5	0.0	0.5	0.9	1.0
2.0	0.06	0.18	0.39	0.72	1
4.0	0.01	0.08	0.25	0.63	1
10.0	0.00	0.01	0.08	0.46	1
$\infty$	0	0	0	0	1

Az oszlopok két perem közti különböző korreláció mellett, a sorok pedig két különböző szabadságfok ( $\nu$ ) paraméter mellett mutatják az aszimptotikus függőség értékeit. Látható, hogy a korreláció növekedésével illetve a szabadságfok csökkenésével nő az aszimptotikus függőség. Mivel  $\nu \rightarrow \infty$  normális eloszlás, ezért a normális eloszlás szimptotikusan független. A 15. ábrán egy gaussi illetve egy  $t_2$  kopulából vett generált mintát ábrázoltunk  $r = 0.8$  korreláció mellett. A  $t_2$  kopulánál jól látható az együttes extrém események nagyobb gyakorisága, a pontfelhő elnyúltabb, több a kiugró érték.

További függőségi struktúrák leírására és függőségek mérésére ad lehetőséget a *kopula* fogalma (ld. 6. definíció). Ennek segítségével adott peremeloszlásokat tetszőleges függőségi struktúrával kapcsolhatjuk össze. Kettőnél több dimenziós esetben pedig speciális konstrukciókkal olyan kopulákat definiálhatunk, melyek a páronkénti függőség mellett rendelkeznek ún. *globális függőségi* paraméterrel.

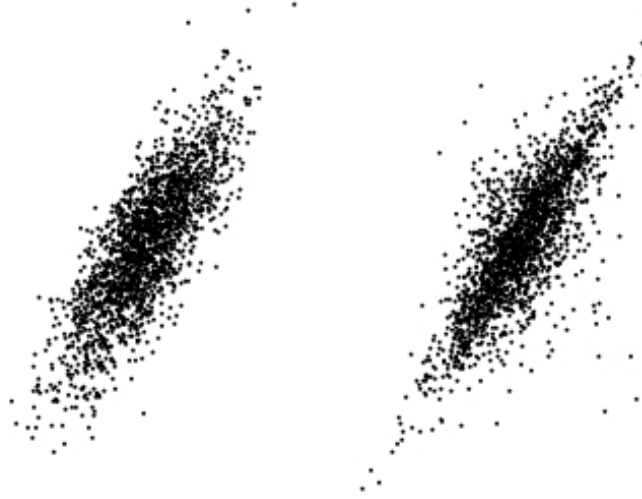
Ebben az alponban olyan modelleket fogunk bemutatni, melyek a normális eloszlásra épülő modellek lehetséges alternatívái. Az aszimptotikus függőség és kopula fogalmak segítségével fogunk modellt építeni, melyet főleg a könnyű utótesztelhetősége miatt *VaR* számításokkal fogunk illusztrálni. Három részvényből álló portfóliókat fogunk megvizsgálni, és pedig az

$$\Pi = \sum_j \pi_j z_j \quad j \in \{1, 2, 3\}$$

alakú egyszerű portfóliókat, ahol  $\pi_j$  az  $j$  indexű papír értéke,  $z_j$  pedig annak hozama. Jelölje  $F_\Pi$  a portfólió eloszlásfüggvényét, ekkor a *VaR* érték kiszámítása  $\alpha$  valószínűségi szinten formálisan

$$VaR_\alpha(\Pi) = -\inf \{v \in \mathbb{R} : F_\Pi(v) \geq 1 - \alpha\}$$

vagyis a portfólió jövőbeli értékeinek az adott valószínűségi szinthez tartozó percentilise. Ha az  $F_\Pi$  eloszlásfüggvény inverze nem ismert, akkor a



15. ábra.

Gaussi és  $t_2$ -kopulából generált minták normális eloszlású peremekkel,  
 $r = 0.8$

$VaR$  értéket Monte-Carlo szimuláció segítségével tudjuk becsülni. A technika pontosan ugyanaz, mint mikor a Shapiro-Wilk statisztikához kritikus értékeket kerestünk. Esetünkben mondjuk  $r$  különböző lehetséges esemény esetén  $\alpha = 0.95$  szinten a  $0.05r$ -dik legrosszabb elem. A papírok értékét egységnyiinek választjuk, vagyis  $\pi_j = 1$  minden  $j - re$ .

Tegyük fel, hogy van a részvények hozamára vonatkozóan egy  $n$  méretű  $Z_j^{(1)}, \dots, Z_j^{(n)}$ ,  $j = 1, 2, 3$  mintánk, és erre szeretnénk modellt illeszteni. Ekkor két lehetőségünk van. Felírhatunk egydimenziós modellt a portfólió eloszlására, vagyis képezzük a  $Z^{(1)}, \dots, Z^{(n)}$  egydimenziós mintát, ahol  $Z^{(i)} = \sum_j Z_j^{(i)}$  és erre illesztjük  $F_\Pi$ -et, majd ennek becsüljük a megadott küszöbértékét. A másik út az, ha egy többdimenziós  $G_3$  eloszlást illesztünk a hozamok többdimenziós mintájából, majd pedig a  $\zeta = \sum_j \zeta_j$  változóra számítunk percentiliseket, ahol  $[\zeta_1, \zeta_2, \zeta_3] \sim G_3$ . Utóbbi bonyolultabb, de kifinomultabb megoldás, hiszen a portfólióban egymástól jelentősen eltérő kockázatosságú instrumentumok is lehetnek, melyeket a portfólió modell esetleg - az aggregáció miatt - összeemos. Mi mindkét megközelítést bemutatjuk.

A  $VaR$ , mint mondtuk, egy jövőbeli értékre vonatkozó becslés. Egy adott minta időszakra, melyet *ablaknak* is szokás nevezni, megbecsüljük az eloszlást. Mivel feltettük, hogy a variancia állandó, ezért egy  $t = 1, \dots, T$  idősort egy mintának tekintünk, és erre illesztjük a kiválasztott eloszlást. A

$VaR_\alpha$  érték a becsült eloszlás  $1 - \alpha$  percentilise, melyet a  $(T + 1)$ -dik napra vonatkoztatunk. A tárgynapra vonatkozó becslésünket tehát a tárgynapot megelőző  $T$  nap eseményeire alapozzuk. A becslés alapján hozott döntésünk bizonyos kockázatot is magában rejt. Egy  $\alpha = 95\%$ -os  $VaR$  érték ugyanis azt jelenti, hogy az esetek 5%-ában tévedni fogunk. Azt, hogy modellünk helyes-e, vagyis hogy ténylegesen ekkora-e a hibaszázalék, úgy ellenőrizhetjük, hogy több becslést végzünk, és megvizsgáljuk a hibaarányt empirikusan. A  $t = 1, \dots, T$  ablakot tovább csúsztatjuk a  $t = 2, \dots, (T + 1)$  időszakra, és megbecsüljük a  $(T + 2)$  tárgynapi  $VaR$  értéket és így tovább. A kapott becslésekből aztán hibaarányt számolunk, és teszteljük (ld. később), hogy az empirikus hibaarány eltérése az elméleti értéktől szignifikáns-e (elfogadható-e), vagy sem.

**Egydimenziós portfólió modellek.** Két modellt vizsgáltunk. Az elsőben a portfólió eloszlására normalitást feltételeztünk, míg a másodikban a portfóliót  $t$ -eloszlásúnak feltételeztük. Mivel a szabadságfok növekedésével a  $t$ -eloszlás a normális eloszláshoz tart, ezért elméletileg csak egy modellről van szó. A gyakorlatban a  $\nu = 50$  szabadsági fok felett a két eloszlás között nincs is különbség. Mivel a  $t$ -eloszlás a  $(\mu, \sigma)$  paramétereiben folytonos, a  $\nu$  paraméterben pedig diszkrét, ezért iteratív becslést használtunk. Adott  $\nu \in [1, 2, \dots, 50]$  értékekre előállítottuk a  $(\mu, \sigma)$  paraméter együttes likelihood becsléseit, majd azt választottuk ki, melynél a likelihood értéke a legnagyobb volt. Mivel  $\nu$ -ben a likelihood értékek „konkáv”-ak, ezért a becslést  $\nu = 1$ -től kiindulva addig érdemes folytatni, míg a likelihood érték el nem kezd csökkenni. Normális eloszlást feltételezve a paraméterek likelihood becslése az egyszerű mintaátlag és a korrigálatlan tapasztalati variancia.

A becsült normális eloszlás megfelelő percentilis értékeit a normális eloszlás inverz eloszlásfüggvényéből (*Wichura [1988]*), a  $t$ -eloszlás percentiliseit pedig az inverz béta-függvény segítségével számíthatjuk ki.

**Kopula modellek.** Két modellcsaládot vizsgáltunk. Az egyikben a függőségi struktúrát  $t$ -kopulával írjuk le, a másikban ún. *MMx* kopulákat illesztünk, melyeket rövidesen ismertetünk. Peremeknek azonban mindkét esetben  $t$ -eloszlást választottunk. A gyakorlatban ugyanis a hozamok eloszlását vagy  $t$ -eloszlásúnak, vagy  $\alpha$ -szimmetrikus eloszlásúnak (ld 32. definíció) feltételezik.

**88. Megjegyzés.** *Ha úgy tetszik, két rivális modellről van szó. Az  $\alpha$ -szimmetrikus (vagy stabil) eloszlás használata elméletileg jobban indokolható, a  $t$ -eloszlás viszont igen jól illeszkedik a hozamok idősoraira. Az érdeklődő olvasónak ajánljuk a Palágyi [2001] dolgozatot, mely egy kollégánk ugyancsak a témában írott disszertációja.*

A kopula modellek becslésének menete a következő:

1. Megbecsüljük a  $t_j(z_j; \mu_j, \sigma_j, \nu_j)$  peremeloszlásokat az előzőekben leírt módon, majd kiszámítjuk a  $U_j^{(i)} = t_j(Z_j^{(i)}; \hat{\mu}_j, \hat{\sigma}_j, \hat{\nu}_j)$   $j = 1, 2, 3$  és  $i = 1, \dots, n$  valószínűségeket.
2. Az így kapott  $U_j^{(n)}, \dots, U_j^{(n)}$   $j = 1, 2, 3$  többdimenziós mintára illesztjük a  $C_3(\mathbf{u}; \theta)$  kopulát szintén likelihood módszerrel.
3. A becslt eloszlás tehát a

$$G_3 \equiv C_3 \left( t_1(z_1; \hat{\mu}_1, \hat{\sigma}_1, \hat{\nu}_1), t_j(z_2; \hat{\mu}_2, \hat{\sigma}_2, \hat{\nu}_2), t_3(z_3; \hat{\mu}_3, \hat{\sigma}_3, \hat{\nu}_3); \hat{\theta} \right)$$

alakot ölti, melyből aztán Monte-Carlo szimulációval  $r$  független véletlen vektort generálunk. Legyen ez a  $\zeta_j^{(1)}, \dots, \zeta_j^{(r)}$ ,  $j = 1, 2, 3$  halmaz.

4. Képezzük a  $\zeta^{(i)} = \sum_j \zeta_j^{(i)}$  statisztikákat, és rendezzük őket, vagyis legyen  $\zeta^{(1)} \leq \dots \leq \zeta^{(r)}$  A  $Var_{\alpha}$  érték az  $r(1 - \alpha)$ -dik legkisebb elem lesz.

**89. Megjegyzés.** A kopula egy eloszlásfüggvény, tehát a paramétereit a szokásos módszerekkel becsüljük. Ha a likelihood becslést választjuk, akkor persze szükség van a sűrűségfüggvényre is.

A  $t$ -kopula a  $t$ -eloszlás eloszlásfüggvényéből az 1.1. pont 8. következménye alapján származtatható. Ha tehát  $H_T(\cdot)$  folytonos  $m$ -dimenziós  $t$ -eloszlás  $t_1(\cdot), \dots, t_m(\cdot)$  peremeloszlásokkal és  $t_1^{-1}(\cdot), \dots, t_m^{-1}(\cdot)$  inverz eloszlásfüggvényekkel, akkor a

$$C_T(u_1, \dots, u_m) = H_T(t_1^{-1}(u_1), \dots, t_m^{-1}(u_m))$$

eloszlásfüggvény a  $t$ -kopula, mely egyértelmű. Ugyanez igaz bármely más folytonos eloszlásfüggvényre is. Ha a kopulát elliptikus eloszlásból származtatjuk a fent leírt módon, akkor *elliptikus kopulának* hívjuk.

Megjegyezzük, hogy különböző peremeloszlásokkal rendelkező eloszlás normális vagy  $t$ -kopulával már nem elliptikus eloszlás, az elliptikusság a függőségi struktúrájára igaz, az eloszlásra magára már nem. A mi modelljeinknél pedig ez a helyzet, hiszen minden egyes peremre külön illesztünk  $t$ -eloszlást, melyek szabadságfokai különbözhetnek. De mi lehet a forma paraméter (illetve korrelációs együttható) jelentése ebben az esetben? Tegyük fel, a korreláció egységnyi a  $k$ -dik és  $j$ -dik perem között, ami lineáris függvényyszerű kapcsolatot jelent, vagyis  $u_j = u_k$ <sup>23</sup>, ahol  $u_j$  és  $u_k$  a kopula  $k$ -dik és  $j$ -dik pereme. Ekkor az eloszlás megfelelő két perem változója  $z_j = t_j^{-1}(u_j)$ ,  $z_k = t_k^{-1}(u_j)$  és így  $z_j = F_j^{-1}(F_k(z_k))$ , vagyis  $z_j$  nemlineáris függvénye  $z_k$ -nak. A válasz tehát, hogy a nemlineáris páronkénti korrelációt méri.

<sup>23</sup>A lineáris függvényyszerű kapcsolat azt jelenti, hogy  $u_j = \alpha u_k + \beta$ . Ugyanakkor a  $t$ -kopula esetén  $\alpha = 1$  és  $\beta = 0$  mivel a  $u_j = t_j^{-1}(x_j)$  változók nulla várható értékű és egységnyi varianciájú véletlen változók.



**90. Megjegyzés.** Az előbbiekben a függőség (és a kockázat) alternatív megközelítéseire koncentráltunk. A kopulák használatával azonban nemcsak a gaussi modellek világából lépünk ki, hanem az elliptikus eloszlásokéból is. Már egy egyszerű  $t$ -kopula modellel is, melyekhez  $t$ -peremeloszlásokat rendelünk, minőségileg nagyot lépünk, hiszen amellet, hogy az egyes instrumentumok viselkedését egyedileg tudjuk modellezni, a lineáris korreláció fogalmát implicate általánosítjuk.

Az elliptikus kopulák származtatása tehát relatíve egyszerű feladat, ebből a világból kilépve azonban többdimenziós kopulák konstruálása nem egyszerű feladat. Általános esetben többdimenziós kopulák létrehozásában fontos szerepet töltenek be a *Laplace-transzformációk*. Jelen alkalmazásban két ilyen transzformációt mutatunk be, az általános tárgyalást illetően ajánljuk a Joe [1997] művet.

**91. Definíció.** *LTA. (pozitív stabil)*

$$\psi(s) = \exp\left(-s^{1/\theta}\right) \quad \theta \geq 1 \quad (31)$$

**92. Definíció.** *LTB. (gamma)*

$$\psi(s) = (1 + s)^{-1/\theta} \quad (32)$$

Ezek segítségével felépíthető a kopulák egy olyan családja, mely igen általános függőségi struktúrával rendelkezik. Mivel a kopulák elmélete igen messzire vezet, ehelyütt nem kívánunk részletekbe bocsátkozni. A (33) formula a kopulák egy ilyen családját definiálja, melynek  $m(m-1)/2$  páronkénti függőségi paramétere van, egy darab globális függőségi paramétere és  $m$  darab szimmetria paramétere van.

$$C(\mathbf{u}) = \psi \left[ - \sum_{i < j} \log K_{ij} \left( e^{-p_i \psi^{-1}(u_i)}, e^{-p_j \psi^{-1}(u_j)} \right) + \sum_{j=1}^m \nu_j p_j \psi^{-1}(u_j) \right] \quad (33)$$

A felírásban  $\psi(s)$  a Laplace-transzformációt jelenti, a  $K_{ij}$  függvények pedig kétdimenziós kopulákat. A Joe [1997] könyv alapján ezt a családot az *MMx* kopuláknak fogjuk a következőkben nevezni.

A  $K_{ij}$  függvények, vagyis a kétdimenziós kopulák képezik a magasabb dimenziószámú kopulák építőköveit. Kétdimenziós kopulák egyszerű struktúrájúak és egyszerűen konstruálhatóak. Előszeretettel használják őket ugyanakkor, mivel még grafikusán ábrázolhatóak. Ugyanakkor kevésbé praktikusak, hiszen a kétdimenziós portfóliók vizsgálatának nincsen gyakorlati jelentősége. Két fontos kétváltozós kopulát ismertetünk, melyek a szakirodalomban széles körben elterjedtek, és amelyeket ebben a tanulmányban is fel fogunk használni.

**93. Definíció.** *B6 család (Gumbel [1960]).*  $1 \leq \delta \leq \infty$  korlátok mellett

$$C(u_1, u_2) = \exp \left( - \left( z_1^\delta + z_2^\delta \right)^{1/\delta} \right) \quad (34)$$

ahol is  $z_j = -\log u_j$ .

**94. Definíció.** *B7 család (Galambos [1975]).*  $1 \leq \delta \leq \infty$  korlátok mellett

$$C(u_1, u_2) = u_1 u_2 \exp \left( \left( z_1^{-\delta} + z_2^{-\delta} \right)^{-1/\delta} \right) \quad (35)$$

ahol is  $z_j = -\log u_j$ .

A 91-94. definíciókra alapozva a következő három, nem-elliptikus kopulát használtuk fel rövid elemzésünkben. Ha  $\psi$  LTA és  $K_{ij}$ -k a B6 családból valók, akkor az MM1 kopulát kapjuk, ha  $\psi$  LTB és  $K_{ij}$ -k a B7 családhoz tartoznak, akkor az MM2 kopulát kapjuk, ha pedig  $\psi$  LTA és  $K_{ij}$ -k B7 családhoz tartoznak, akkor MM3 kopulát kapjuk meg. A paraméterek értelmezése a következő: A  $\psi(s)$ -hez tartozó  $\theta$  paraméter határozza meg a globális függőség mértékét, a  $K_{ij}$  kopulák  $\delta_{ij}$ oefficiensei a páronkénti függőségeket,  $\nu_j$  paraméterek pedig a szimmetria mértékét határozzák meg. Tekintve, hogy behelyettesítés után a konkrét formák igen bonyolultak, ezért ehelyütt nem jelenítjük meg. Az érdeklődő olvasó megtalálja a levezetéseket a forráscikkben.

**95. Megjegyzés.** *Bizonyítható, hogy a felsorolt MM1-MM3 kopulák mindegyike aszimptotikusan összefüggő. Emellett láttuk, hogy ennek a kopulacsaládnak van globális függőségi paramétere is ( $\theta$ ). Mint elemzéseinkből majd kiderül, a  $t$ -kopulánál hasonló funkciót tölt be a szabadságfok paraméter ( $\nu$ ). Magasabb szabadságfok ugyanis kisebb aszimptotikus és kisebb globális függőséget jelent és megfordítva.*

Mivel a modell becslését likelihood módszerrel fogjuk végezni, ezért szükségünk van a sűrűségfüggvényekre. Ez, mint ismeretes, az eloszlásfüggvény  $m$ -rendű vegyes parciálisa, vagyis

$$c_{MMx}(\mathbf{u}) = \frac{\partial^m C_{MMx}(\mathbf{u})}{\partial u_1 \dots \partial u_m}$$

A (33) képletből látható, hogy az

$$h(x_1, \dots, x_m) = f[g(x_1, \dots, x_m)]$$

általános formulának kell a vegyes parciálisait előállítanunk  $m$  foksámgig. Mivel a  $g(\cdot)$  leképzés a  $K_{ij}$  kétdimenziós kopulák összegéből áll össze, ezért csak az első- és másodrendű vegyes parciális deriváltjai lehetnek zérustól különbözőek, ami némileg egyszerűsíti a számításokat.

**96. Megjegyzés.** Mi egy egyszerű szimbolikus deriváló algoritmust készítettünk, mely az MMx kopulák vegyes parciálisait előállítja tetszőleges  $m$ -re. Megjegyezzük, hogy  $m = 6$  esetén 96 tagú a derivált, mely igen nagy ütemben emelkedik  $m$  növelésével. Az MMx sűrűségfüggvényének előállítása tehát számítástechnikai szempontból igen költséges.

A Student  $t$ -kopulából való mintavétel viszonylag egyszerű feladat. Első lépcsőben egy  $x_1, \dots, x_m$  véletlen mintát veszünk az  $(\mathbf{R}, \nu)$  paraméterű többdimenziós  $t$ -eloszlásból. Ehhez először a többdimenziós standard  $t$ -eloszlásból veszünk egy véletlent vektort, mégpedig a

$$\mathbf{x} = \frac{\mathbf{z}}{\chi_\nu / \sqrt{\nu}} \quad \mathbf{z} \sim N_m(\mathbf{0}, \mathbf{I}), \quad \chi_\nu = \sqrt{\xi_1^2 + \dots + \xi_\nu^2}, \quad \xi_j \sim i.i.d. N(0, 1)$$

formula alapján, majd a (27) definíció alapján a kívánt paraméterekkel rendelkező eloszlású vektorra alakítjuk. Ezt a vektort aztán komponensenként egyenletes eloszlású változókká transzformáljuk a  $u_1 = t_\nu(x_1), \dots, u_m = t_\nu(x_m)$  formula alapján, ahol  $t_\nu(\cdot)$  a standard  $t$ -eloszlást jelenti.

Az MMx kopulák esetén ilyen egyszerű módszer nem áll rendelkezésre, ezért egy másik utat kell követnünk.

- Vegyünk egy  $p_1$  egyenletes eloszlású változót és oldjuk meg a  $p_1 = C(u_1, 1, \dots, 1)$  egyenletet  $u_1$ -re
- Vegyünk egy  $p_j$  egyenletes eloszlású változót és oldjuk meg a  $p_j = C_j(u_j | u_1, \dots, u_{j-1})$  egyenletet  $u_j$ -re  
ahol

$$\begin{aligned} C_j(u_j | u_1, \dots, u_{j-1}) &= \Pr(U_j \leq u_j | U_1 = u_1, \dots, U_{j-1} = u_{j-1}) \\ &= \frac{\partial^{j-1} C(u_1, \dots, u_{j-1}, u_j, \dots, 1)}{\partial u_1 \dots \partial u_{j-1}} \bigg/ \frac{\partial^{j-1} C(u_1, \dots, u_{j-1}, 1, \dots, 1)}{\partial u_1 \dots \partial u_{j-1}} \end{aligned}$$

és ezt ismételjük  $j = 2, \dots, m$ -re.

Ha  $C_j^{-1}(\cdot)$  nem írható fel zárt alakban, akkor  $u_j$ -ket numerikus gyökkereső módszerrel kell kiszámítanunk. A  $C_j(\cdot)$  deriváltjainak kiszámítása az MMx kopulák esetében bonyolult, így célszerű olyan algoritmus használata, mely nem igényli a gradiens kiszámítását. Mi *Brent* módszerét használtuk, mely a *Press et al. [1992]* műben megtalálható.

Mindkét eloszlástípus esetén a mintavétel első lépése mindig az, hogy egyenletes eloszlásból kell egy véletlen vektort generálni, majd ezt transzformáljuk a kívánt eloszlássá. Mivel több időszakra vonatkozóan illesztjük ugyanazokat a modelleket, ezért célszerű a kiinduló egyenletes eloszlásból származó véletlen számoknak ugyanazzal a halmazával dolgozni minden időszakban, és az adott időszakra becsült kopula paraméterek alapján a

kívánt eloszlásba transzformálni. Ezzel csökkenthető a becslések varianciája, hiszen a becslések nincsenek mintavételi ingadozásnak kitéve, csupán a paraméter becslések hibájával terheltek.

A  $t$ -kopula modell paraméterének likelihood becsléséhez lokális optimalizáló módszert használtunk, amivel kedvező tapasztalataink voltak. Az  $MMx$  kopuláknál viszont a feladat komplexitásából adódóan globális módszer használata vált szükségessé. Mi a *Gablonksy [2000]* által készített és publikált Lipschitz optimalizáló algoritmust használtuk fel.<sup>24</sup>

**97. Megjegyzés.** *Elemzésünkben nem mutatunk be többdimenziós gaussi modellt, ugyanis mint látni fogjuk, nem érdemes. A portfóliók egydimenziós eloszlására illesztettünk viszont normális eloszlást. Ha ugyanis a normalitást valószínűségi változók összegének eloszlására követeljük meg, nem pedig minden egyes változóra, akkor a normalitás szempontjából nyilvánvalóan egy engedékenyebb modellt kapunk. Amint látni fogjuk majd, ez a modell is alulmarad a többi alternatívával szemben.*

Két piacot fogunk vizsgálni, az amerikai és a magyar piacot. A 3-3 *Dow Jones* és *BUX* részvényből készítünk portfóliókat, és ezek hozamaira próbáljuk meg a fent leírt 6 modellt illeszteni. A modellek értékelését három kritérium alapján végezzük. Először  $\chi^2$ -teszt segítségével a modellek illeszkedését vizsgáljuk meg. Ezután  $\alpha = 95, 99$  és  $99.5\%$ -os *Var* értékeket számolunk, melyeket szintén formális teszt alapján értékelünk. A formális tesztek után kvalitatív elemzést végzünk, vagyis grafikonok segítségével próbáljuk értelmezni a eredményeinket.

A modellek becslése és értékelése után megvizsgáljuk többdimenziós normalitás tesztjeinket. Az opcióárazásos alkalmazásnál bemutatott, bootstrap technikával végzett elemzést megismételjük a kiválasztott amerikai és magyar részvények hozamainak idősor vektoraira.

### 6.2.1. Dow Jones részvények

Alapadatok:

Minta időszak: 1990. február 26. - 2001. február 22. (2778 megfigyelés)

Részvények: *General Electric (GE)*, *General Motors (GM)*, *Citigroup (CI)*

Ablak mérete: 500 (összesen 2279 periódus)

Monte-Carlo szimuláció mérete:  $r = 10000$

**Illeszkedés vizsgálat.** A portfólió modellek esetében jó illeszkedést tapasztaltunk. Az egydimenziós  $t$ -modell illeszkedését legalább  $5\%$ -os szignifikancia szinten a legtöbb periódusban elfogadtuk. A többdimenziós modellek

<sup>24</sup>A témáról kiváló áttekintést talál az olvasó a *Pintér [1996]* és *Zhigljavsky [1991]* művekben.

köziül a  $t$ -kopula modell és az  $MM2$  modell (mindkettő  $t$ -eloszlású peremekkel) bizonyult szignifikánsnak a tesztek alapján, a legtöbb periódusban az illeszkedés 5%-nál magasabb szinten elfogadható volt. 1996 és 1998 között voltak „gyenge” periódusok, amikor az illeszkedések szignifikanciája alacsonyabb volt, de az 1%-os szint felett maradt. Az  $MM1$  és  $MM3$  kopula modellek a legtöbb esetben még az 1%-os szintű illeszkedést sem érték el.

**VaR becslések.** A formális tesztek eredményeit a 10-15. táblázatokban foglaltuk össze.  $\alpha = 5\%$ ,  $1\%$  és  $0.5\%$  szintű  $VaR$  értéket számoltunk. Az empirikus hibák értékét félkövér betűkkel szedtük és a táblázatok második soraiban találhatóak. Az utolsó sorok tartalmazzák a *Kupiec-teszt* (Kupiec [1995]) értékét illetve az empirikus szignifikanciáját, vagyis a  $p$ -értéket. A teszt alapelve az, hogy a pl. 5%-os  $VaR$  értéknél csak az esetek nagyjából 5%-ában kaphatunk nagyobb veszteséget. A *Kupiec-teszt* azt ellenőrzi, hogy az ettől való eltérés szignifikáns-e vagy csak a mintavételi ingadozás okozza.

10. táblázat. Egydimenziós gaussi modell

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>4.70%</b>	<b>1.32%</b>	<b>1.05%</b>
Kupiec	0.46	2.10	10.61
$p$ -érték	49.99%	14.77%	0.11%

11. táblázat. Egydimenziós Student  $t$ -modell

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>4.91%</b>	<b>1.10%</b>	<b>0.70%</b>
Kupiec	0.04	0.21	1.66
$p$ -érték	85.09%	64.69%	19.75%

Mindkét portfólió modell esetében elfogadható illeszkedést tapasztaltunk, bár a  $t$ -modellre vonatkozó eredmények jóval meggyőzőbbek. A gaussi modellből számított empirikus hibák 1 és 5%-os szinteken meglehetősen nagyok. Ha számításba vesszük, hogy a  $t$ -modell nem sokkal bonyolultabb, és nem igényel jelentősebb számítástechnikai kapacitásokat, mint a normális modell, akkor mindenképpen a  $t$ -modellt érdemes választani.

Az illeszkedés vizsgálat eredményeivel összhangban a  $t$ -kopula és az  $MM2$  modellekből számított  $VaR$  értékek a formális teszten átmertek, míg a másik két modellből számítottak nem. Az  $MM1$  és  $MM3$  kopula modellekből számított empirikus hibák igen magasak, ennek megfelelően a  $p$ -értékek nullák. A  $t$ -kopula modell  $VaR_{0.99}$  értéke kicsit jobb teszt eredményt mutat, mint az  $MM2$  modellből számított, míg  $VaR_{0.95}$  érték utóbbinál bizonyult biztosabbnak. A  $VaR_{0.995}$  értéknél gyakorlatilag azonosak a teszt

12. táblázat. Student  $t$ –kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>4.17%</b>	<b>1.05%</b>	<b>0.61%</b>
Kupiec	3.51	0.06	0.56
$p$ -érték	6.11%	80.06%	45.52%

13. táblázat.  $MM1$  kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>7.15%</b>	<b>2.15%</b>	<b>1.18%</b>
Kupiec	19.72	22.90	15.48
$p$ -érték	0.00%	0.00%	0.01%

14. táblázat.  $MM2$  kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>5.66%</b>	<b>1.10%</b>	<b>0.61%</b>
Kupiec	2.01	0.21	0.56
$p$ -érték	15.62%	64.69%	45.52%

15. táblázat.  $MM3$  kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>7.06%</b>	<b>2.15%</b>	<b>1.23%</b>
Kupiec	18.23	22.90	17.26
$p$ -érték	0.00%	0.00%	0.00%

eredmények.

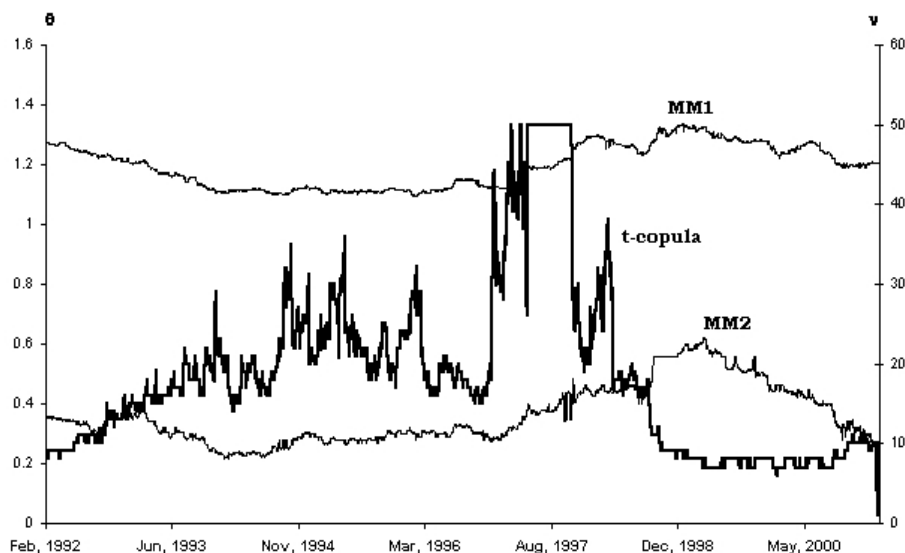
**Kvalitatív vizsgálat.** A *Dow Jones* piacon végbemenő folyamatok jól nyomon követhetőek a globális függőséget mérő paraméterek változásán keresztül. Az *MMx* kopuláknak van is egy ilyen paramétere ( $\theta$ ), míg a *t*-kopula esetében a globális függőséget a szabadságfok paraméteren ( $\nu$ ) keresztül tudjuk megragadni. A 16. grafikonon ezek értékét ábrázoltuk az egyes periódusokra. Mivel az *MM1* és *MM3* kopulák általában hasonló eredményeket produkáltak, így csak egyikőjük paraméterét tettük fel a grafikonra.

Amint az illeszkedés tesztek eredményeiben bekövetkező bizonytalanság is utalt rá, az 1996-1998 közötti időszakban egy drámai változás következett be a piacon, legalábbis a három vizsgált részvényt illetően. Ezekben a periódusokban a globális függőség csökkent (ami a *t*-kopula modell esetében a  $\nu$  paraméter növekedésével jár együtt), a piac úgymond „szétszakadt”. A peremeloszlásokra pillantva azt látjuk (18. ábra), hogy *GE* részvény hozamára illesztett *t*-eloszlás szabadságfoka a  $\nu_{GE} = 30-50$  tartományban mozog, vagyis közelítőleg normális eloszlású. Ezzel ellentétben *GM* hozamoknál  $\nu_{GM} = 7-12$  szabadságfokokat becsültünk, ami leptokurtikus eloszlásra utal. A harmadik részvényre becsült szabadságfokok (*CI*) a  $\nu_{CI} = 33$  értékről  $\nu_{CI} = 7$  értékre csökkentek, az eloszlás 1998 végére csúcsossá, hosszan elnyúlóvá vált<sup>25</sup>. Később a piac ismét visszaállt, és mindhárom részvényre becsült paraméterek visszatértek a korábbi  $\nu = 4-15$  körüli szintre. Az 1996-os eseményekhez hasonló történt újra 2000 tavasza körül. Látszólag a *GE* részvénynek tulajdoníthatók túlnyomó részben az említett változások, hiszen ezen részvény szabadságfoka távolodik el a többiekétől mindkét esetben, ami a globális függőség csökkenését vonja maga után.

A 17. ábrára pillantva azt találjuk, hogy a  $VaR_{0.99}$  értékek pontosan követik a globális függőség változását. Amikor ugyanis az egyes részvények elkezdenek különbözőképpen viselkedni, vagyis diverzifikálódik a portfóliónk, akkor az egymástól való globális függőség csökken, ami alacsonyabb kockázatot jelent, s ezért az egyes modellekből számított *VaR* értékek is alacsonyabbak lesznek.

A teljesség kedvéért egy pár szót szólunk a páronkénti függőségről is. Az *MM2* kopula  $\delta_{ij}$  paraméterének időszora bár zajos, mégis egy struktúra látszik körvonalazódni benne. A  $\delta_{GE,GM}$  paraméter monoton csökkent, míg a  $\delta_{GE,CI}$  paraméter monoton nőtt (vagyis egyre erősebben együtt mozogtak) a teljes vizsgált időszakra a gyenge időszakot kivéve, amikor is a tendencia megfordult. Ezalatt  $\delta_{GM,CI}$  konstans maradt, és gyenge függőséget mutatott. Mivel a páronkénti függőség és a portfólió kockázata között nehéz a kapcsolatot megtalálni, így ezek a paraméterek önmagukban nem sok

<sup>25</sup>Ez a tendencia egészen 2000. augusztusig folytatódott, mikor a  $\nu = 4$  minimális értéket ért el.



16. ábra. Globális függőség - Dow Jones részvények

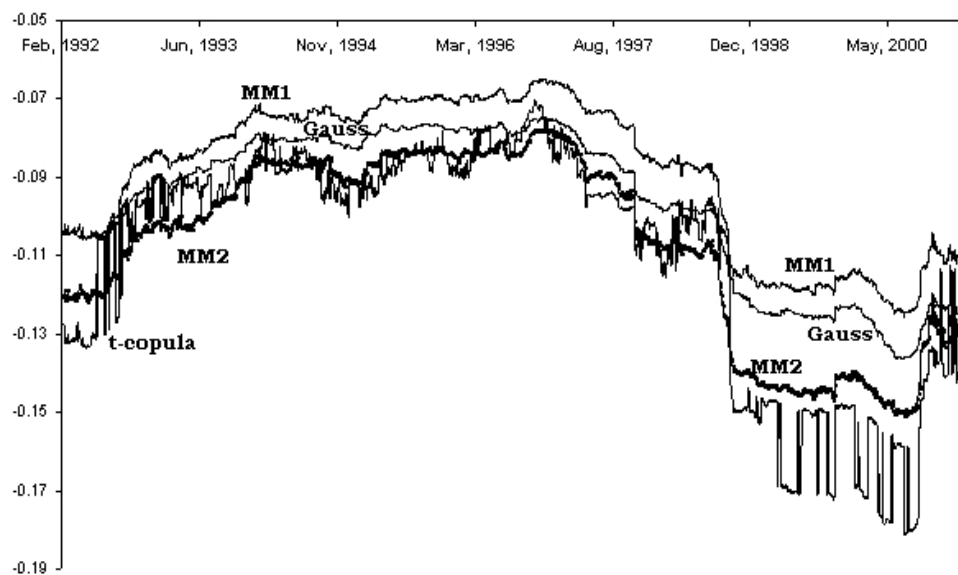
konklúzióval szolgálnak.

Habár a  $t$ -kopula és az  $MM2$ -kopula modellekből nagyjából ugyanakkora  $VaR$  értékeket számolunk, a  $t$ -kopula modell becslései meglehetősen zajosak. Ez különösen a gyenge periódusokban érzékelhető, amikor a  $GE$  hozamok és a kopula szabadságfok paraméterei is magasak. Ez abból a tényből ered, hogy a szabadságfokok magasabb tartományában, például egy  $t_{20}$  és  $t_{25}$  modell között, nincs igazán nagy különbség, míg mondjuk egy  $t_2$  és  $t_3$  modell között igen éles a határ.

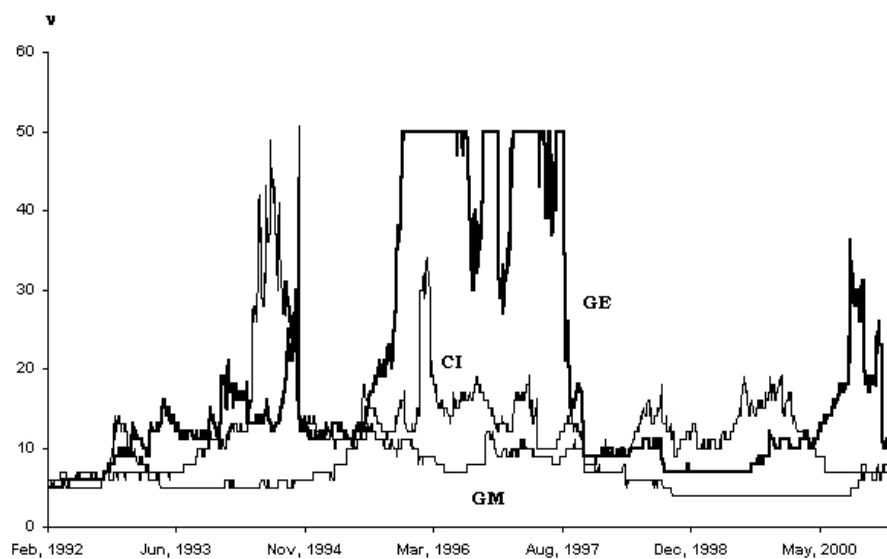
A többdimenziós normalitás tesztek eredményei egyhangúan azt sugallták, hogy a normalitás hipotézise az 1997 január - 1997 április és 2000 november - 2001 január időszakokat kivéve nem állja meg a helyét. Nem meglepően mindkét időszak a gyenge időszakokba esik, amikor is diverzifikáltabb portfóliónk van, és alacsonyabb veszteséget várunk a becsült  $VaR$  értékek alapján. Habár erősen megkérdőjelezhető, hogy ezekben az időszakokban az együttes eloszlás valóban normális<sup>26</sup>, ez részben mégis magyarázatul szolgál arra, hogy a korábbi, a normalitás hipotézisére épített modellek miért látszottak működni bizonyos időszakokban. Ezek a korai modellek akkor valottak kudarcot, mikor extrém események egyszerre következtek be, melyek modellezésére a jelen alponthban bemutatott modellek alkalmasak, a gaussi

<sup>26</sup>A peremeloszlások ugyanis igen eltérőek, ami egy nem szimmetrikus együttes eloszlásra utal. Az egyik perem közel normális eloszlású, míg más peremek leptokurtikus eloszlások, ezek közt még  $t_7$ -eloszlást is találunk.

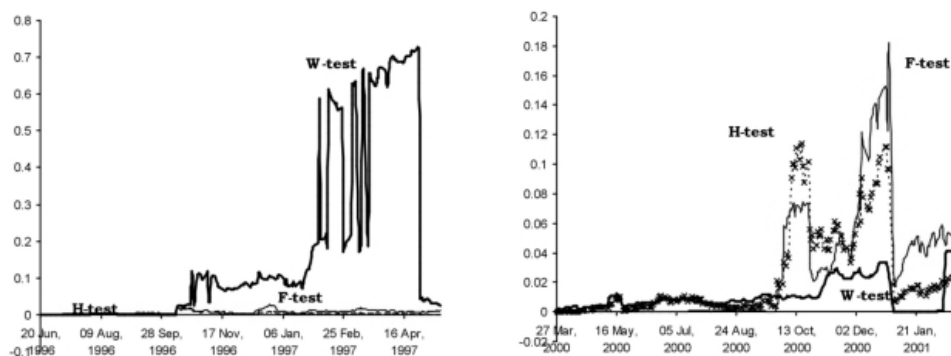




17. ábra.  $VaR_{0.99}$  becslések - Dow Jones részvények



18. ábra. Peremeloszlások - Dow Jones részvények



19. ábra. A többdimenziós normalitás tesztek  $p$ -értékei

modellek nem, és ezért többnyire jelentős mértékben alulbecsülik a várható veszteségeket.

### 6.2.2. BUX részvények

Alapadatok:

Minta időszak: 1997. november 17. - 2001. május 8. (864 megfigyelés)

Részvények: *MOL Rt.*, *Matáv Rt.*, *OTP Bank Rt.*

Ablak mérete: 250 (összesen 615 periódus)

Monte-Carlo szimuláció mérete:  $r = 10000$

**Illeszkedés vizsgálat.** A portfólió modellek, különösképpen az egydimenziós  $t$ -modell, igen jó illeszkedést produkáltak, a tesztből számított empirikus szignifikancia szintek 5% felett voltak a legtöbb időszakban. A  $t$ -kopula modell hasonlóan jól illeszkedett 5% illetve 10% szinten a legtöbb periódusban. Az  $MMx$  modellek egyike sem volt szignifikáns a teljes vizsgált időszakban. Az  $MM2$  jól illeszkedett az első periódusokban, 5%-os szinten elfogadható volt, majd 2000. november környékén már 1% alá esett. Az  $MM1$  és  $MM3$  modellek az 1999. szeptember és 2000. november közti periódusokban mutatott szignifikáns illeszkedést. Ezután az időpont után csak a  $t$ -kopula modell illeszkedett, és az illeszkedése nem romlott.

**VaR becslések.** A formális tesztek eredményeit a 16-21. táblázatokban foglaltuk össze.

Az egydimenziós normális és  $t$ -modellekre hasonló eredményeket kaptunk. A  $VaR_{0.95}$  ( $\alpha = 5\%$ ) becslések meglehetősen konzervatívak, ami az alacsony backtest értékekben tükröződik. A  $VaR_{0.99}$  és  $VaR_{0.995}$  veszteségeket a  $t$ -modell felül-, míg a normális modell alulbecsüli. Az összes

16. táblázat. Egydimenziós gaussi model

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>3.58%</b>	<b>1.14%</b>	<b>0.81%</b>
Kupiec	2.90	0.11	1.02
$p$ -érték	8.87%	73.61%	31.31%

17. táblázat. Egydimenziós Student  $t$ –modell

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>3.74%</b>	<b>0.81%</b>	<b>0.33%</b>
Kupiec	2.24	0.23	0.43
$p$ -érték	13.41%	63.00%	51.14%

számított becslés szignifikáns 5%-os szinten. Mindent összevetve a  $t$ -modell a jobb alternatíva.

18. táblázat. Student  $t$ –kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>1.79%</b>	<b>0.65%</b>	<b>0.33%</b>
Kupiec	17.54	0.87	0.43
$p$ -érték	0.00%	35.20%	51.14%

Az illeszkedés vizsgálat által hozott eredményekkel ellentétben az  $MM1$  és  $MM3$  kopula modellek igen jó becsléseket adtak. Az empirikus hibák közel estek a megadott  $\alpha = 5, 1$  és 0.5%-os értékekhez és szignifikánsnak bizonyultak. A  $t$ -kopula és az  $MM2$ -kopula modellek becslései meglehetősen konzervatívak, de minden szignifikancia szinten elfogadhatóak (kivéve a  $t$ -kopula modell  $VaR_{0.95}$  becslését). Megjegyezzük, hogy mivel az  $MM2$  az illeszkedés teszten nem mindenhol ment át, ezért teljes biztonsággal csak a Student  $t$ -kopula modell ajánlható.

**Kvalitatív vizsgálat.** Mivel a minta időszak meglehetősen rövid volt, ezért igazán komoly következtetéseket nem lehet az eredményekből levonni. Ez a piac egészen más mögöttes logika szerint működik, mint a *Dow Jones* részvények piaca. Valójában ez a piac sokkal érdekesebb is, hiszen az extrém események együttes bekövetkezésének valószínűsége sokkal nagyobb, ezért a befektetések várható kockázata is magasabb.

Először pillantsunk a 20. grafikonra, ahol a globális függőséget meghatározó paraméterek idősorát ábrázoltuk a  $t$ -kopula, az  $MM1$  és  $MM2$  modellek esetében. A helyzet a következő: Nagyjából 1999. decemberéig egy kockázatos időszakot találunk, amit a  $t$ -kopula modellre becsült alacsony,

19. táblázat. *MM1* kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>4.39%</b>	<b>1.14%</b>	<b>0.49%</b>
Kupiec	0.50	0.11	0.00
<i>p</i> -érték	47.90%	73.61%	96.57%

20. táblázat. *MM2* kopula

VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>3.90%</b>	<b>0.65%</b>	<b>0.33%</b>
Kupiec	1.68	0.87	0.43
<i>p</i> -érték	19.47%	35.20%	51.14%

$\nu = 4 - 7$  szabadságfokok, illetve az *MMx* modellekre becsült magas  $\theta$ -ák sejtetnek. Megjegyezzük, hogy *MM1* és *MM2* modelleket különböző *LT*-vel konstruáltuk, így  $\theta$ -iknak is más az alsó korlátja. Egy alacsony globális függőséggel jellemzett csendesebb időszak következik nagyjából 2000. november-decemberéig, azután pedig a globális függőség ismét növekedni kezd. Ebben az utolsó szakaszban gyengül az *MMx* modellek illeszkedése.

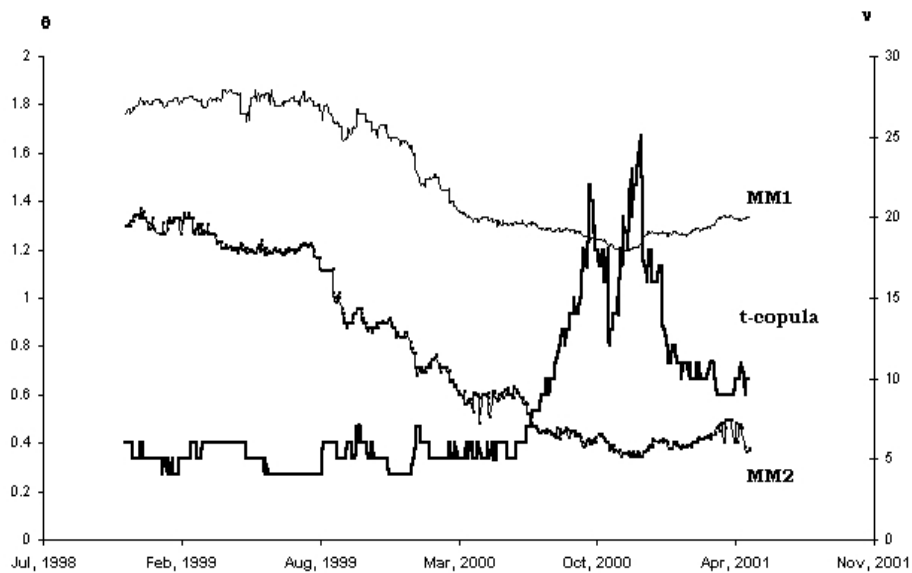
A 99%-os *VaR* értékek ugyanezt a mozgást írják le, amint azt a 21. ábrán láthatjuk. A kockázatos időszakban a várható veszteségek magasabbak, míg az 1999. decemberétől 2000. decemberéig tartó időszakban ez mérséklődik. Ezután pedig a portfólió ismét kockázatosává válik.

Ha megnézzük a peremeloszlásokat (ld. 22. grafikon), világosan látszik, hogy egyetlen részvény (*MOL*) egymagában felelős a portfólió diverzifikációjáért. 1999 decemberéig a portfóliónkban  $\nu = 4 - 7$  szabadságfokú *t*-kopula és  $\nu = 4 - 5$  szabadságfokú *t*-peremeloszlásokkal leírható hozamokkal bíró részvények vannak. Ettől az időponttól kezdve viszont a *MOL* papírok hozamának eloszlása kezd a normális eloszláshoz közelíteni (vagyis a szabadságfoka növekedni kezd, míg nem eléri a maximális  $\nu = 50$  értéket). Ezzel párhuzamosan csökken a globális függőség, a kopula szabadságfoka  $\nu = 25$  értékig emelkedik. Mikor pedig  $\nu_{MOL}$  ismét csökkenni kezd, a globális függőség és a portfólió kockázatosága megint növekedni kezd.

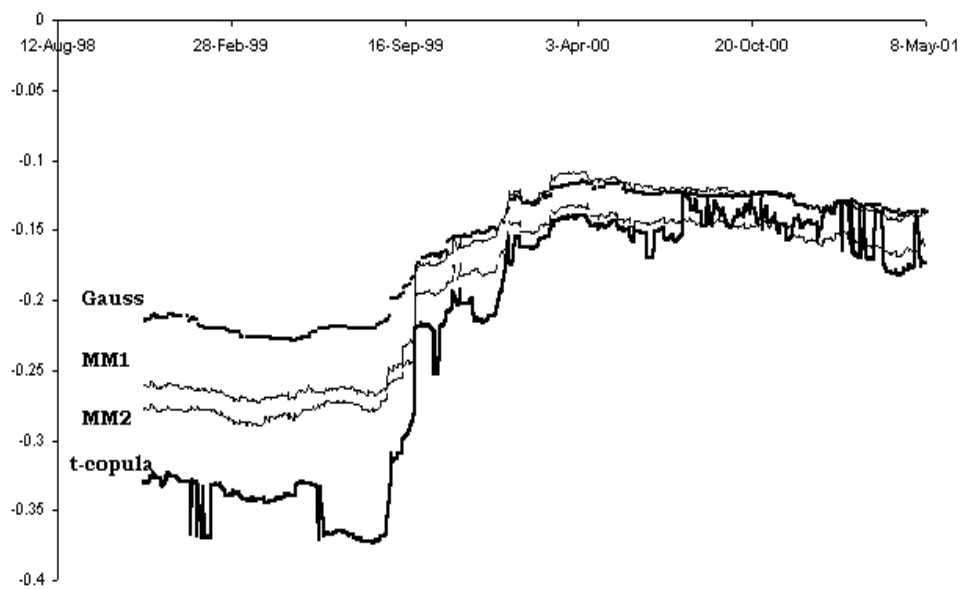
A többdimenziós normalitás tesztek az együttes normalitásnak semmi-

21. táblázat. *MM3* kopula

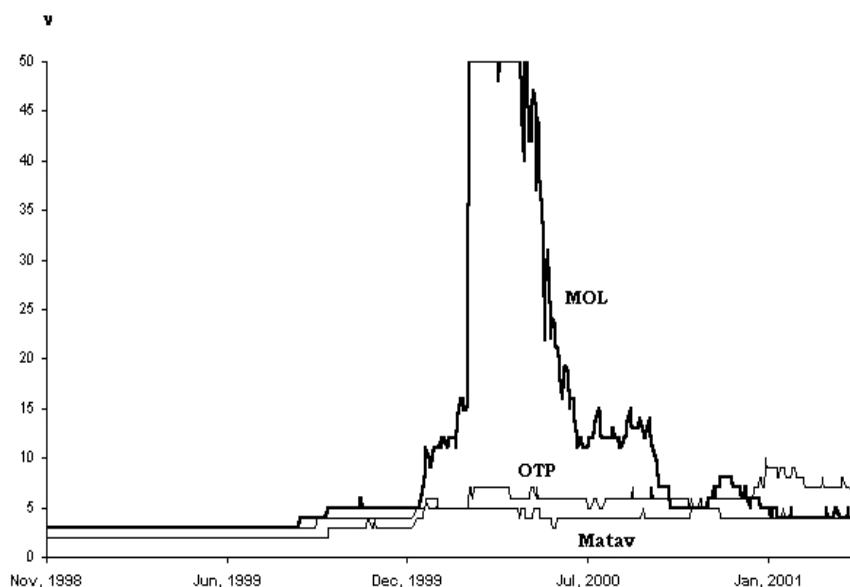
VaR konfidencia szint	5%	1%	0.5%
Hibaarány	<b>4.23%</b>	<b>1.14%</b>	<b>0.49%</b>
Kupiec	0.81	0.11	0.00
<i>p</i> -érték	36.72%	73.61%	96.57%



20. ábra. Globális függőség - *BUX* részvények



21. ábra.  $VaR_{0.99}$  becslések - *BUX* részvények



22. ábra. Peremeloszlások - *BUX* részvények

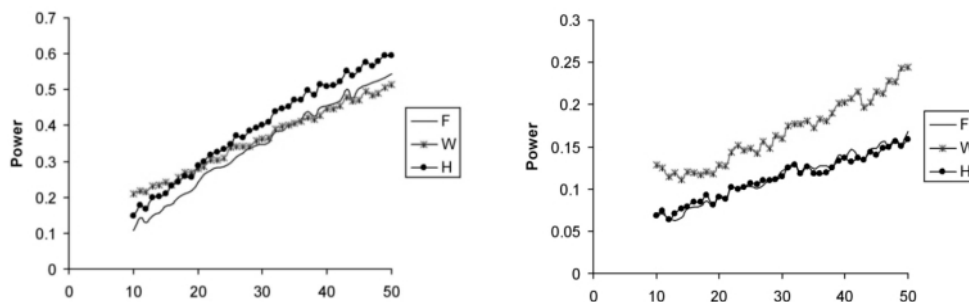
lyen jelét sem mutatták. A magyar piac kockázatossága miatt tehát egy tipikusan olyan alkalmazási terület, ahol a normalitásra épülő modelleknek nincsen gyakorlati relevanciája.

### 6.2.3. A próbák erejének vizsgálata

Az előbbi rövid elemzések alapján arra a következtetésre juthatunk, hogy a  $t$ -kopula és az  $MM2$ -kopula modellek  $t$ -eloszlású peremekkel jól jellemzik a *Dow Jones* részvények hozamainak idősorait. Az illeszkedés vizsgálatok szignifikáns illeszkedést mutattak, és a becsült  $VaR$  értékek is átmentek a formális teszten. A portfólió modellek is elfogadhatóan illeszkedtek, de a  $VaR$  értékek tesztje bizonytalanabb volt, és mint a grafikonon láttuk, a gaussi modell által becsült veszteségek sokkal alacsonyabbak voltak, mint az előbbi két modell esetében. Az előrejelzett veszteségek közti különbség abban az időszakban volt a legkisebb, mikor a többdimenziós normalitás tesztek normalitást véltek az idősor vektorban felfedezni. A kockázatosabb időszakokban azonban a  $VaR$  olló kinyílt, s ez a *BUX* részvények esetében - lévén magasabb kockázatúak - sokkal szembetűnőbb volt.

A *BUX* részvények piacán is hasonló következtetésekre jutottunk a modellválasztással kapcsolatban. Habár a Kupiec-teszt a teljes időszakra vonatkozóan elfogadhatónak ítélte az egydimenziós normális eloszlású modellt, mégis sokkal kisebb veszteségeket becsültünk belőle<sup>27</sup>, mint a többi mo-

<sup>27</sup>A backteszt hibák szintén alacsony  $VaR$  értéket mutattak, de a kevésbé kockázatos



23. ábra. Próbák ereje, *Dow Jones* részvények hozamai

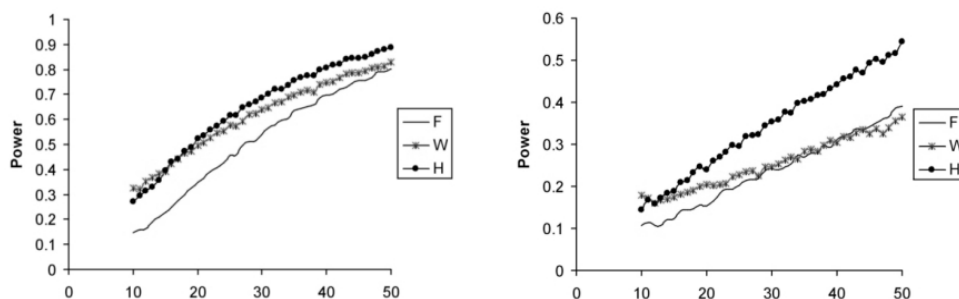
dellekből a kockázatosabb időszakban (ld. 21. ábra). Ezzel ellentétben a  $t$ -kopula modell rendkívül konzervatív becsléseket eredményezett, vagyis túlbecsülte a várható veszteségeket, míg az  $MM2$ -kopula valahol a kettő között foglal helyet.

Az idősor vektorokra vonatkozóan elvégeztük a normalitás teszteket, ebben az alponthoz megvizsgáljuk a próbák erejét kis mintákra *bootstrap* módszerrel. Mind a *Dow Jones*, mind pedig a *BUX* papírok esetében két-két ablakot választottunk ki: egyet a kockázatosabb, egyet pedig a kevésbé kockázatos időszakból. Az ablakból  $n = 10, \dots, 50$  méretű mintákat vettünk, mégpedig  $r = 5000$  darabot, és kiszámítottuk mindegyikre az  $F$ ,  $W$  és  $H$ -teszteket. Számláltuk, hogy az  $r$  darab mintából hányat utasítanak el, az elutasítási hányaddal becsültük a próbák erejét. Eredményeinket a 23-24. ábrákon mutatjuk be. A 23. ábra bal oldali grafikonja az amerikai piac egy kockázatosabb időszakában mutatja a hozameloszlásra vonatkozó vizsgálatot, a jobb oldali pedig egy nyugodtabb időszakban. A kockázatos időszakban a tesztek közt nincs jelentős különbség, bár a  $H$ - és  $W$ -teszt némileg erősebb. A mintaméret növekedésével a próbák ereje növekszik.  $n = 40$  méretű mintánál a próbák ereje  $0.4 - 0.5$ ,  $n = 50$  mintánál pedig  $0.5 - 0.6$ . A kevésbé kockázatos időszakban a próbák ereje kisebb, az  $F$  és  $H$ -teszt esetében  $n = 50$  méretű mintánál mindössze  $0.15$ . A  $W$ -teszt ugyanakkor jóval erősebbnek látszik a többinél, ami igen pozitív dolog, hiszen a normalitás hiányát akkor kell nagy biztonsággal felismerni, mikor az nem nyilvánvaló.

A magyar részvények esetében (24. ábra) a próbák ereje valamelyest nagyobb. A *BUX* részvények hozamainak eloszlására ugyanis a csúcsosság és a magasabb globális függőség (alacsonyabb szabadsági fok a  $t$ -modelleknél) jellemző, ezért a tesztek biztosabban felismerik a normalitás hiányát. A kevésbé kockázatos időszakban azonban most a  $H$ -teszt látszik a legerősebbnek, míg az  $F$  és  $W$ -tesztek hasonlóak. Mivel a  $H$ -teszt elsősorban a perem

---

periódusok kompenzálták az itt vétett hibákat.



24. ábra. Próbák ereje, *BUX* részvények hozamai

információkra épít, ezért vélhetően igen érzékeny a magyar papíroknál tapasztalható alacsony szabadságfokú  $t$ -peremekre. Ez azt sugallja, hogy az ilyen típusú tesztekre is szükség lehet, hiszen előfordulhat, hogy a peremekre vonatkozó információ árulkodóbb, mint az eloszlás többdimenziós struktúrája.

### 6.3. Következtetések

A Shapiro-Wilk teszt különböző variánsait vizsgáltuk meg valós adatsorokon. Az első alkalmazásnál a peremek vizsgálata alapján a többdimenziós normalitás hipotézisét nem lehetett elvetni, ezért az együttes normalitás tesztelése kulcsfontosságú volt. Nem meglepő módon a  $H$ -teszt gyengének bizonyult ezekre az adatsorokra, hiszen ez a teszt a perem információkat kombinálja. A  $K$  paraméter növekedésével azonban az egyik perem egyre inkább „eldeformálódott“, melyre már a  $H$ -teszt is reagált. A legjobb teljesítményt a  $W$ -teszt nyújtotta, és a normalitás vizsgálatára vonatkozó eredmények a torzítás vizsgálat eredményeivel is konzisztensek voltak. A torzítás ott vált zérussá, ahol a kismintás bootstrap elemzésekben a próbák empirikus értéke a legnagyobb volt.

A második alkalmazásnál hozamokra vonatkozó adatsor vektoraink voltak, melyről tudjuk, hogy jellemzően csúcsos, hosszan elnyúló eloszlások. Emellett a modell illesztésnél azt tapasztaltuk, hogy a permeloszlások gyakran igen eltérőek. Ezen két jellemző nagy biztonsággal kizárja az együttes normalitást, mégis voltak olyan ablakok, mikor ennek jelét láttuk. Ezzel különös összhangban a becült veszteségek is itt voltak a legalacsonyabbak. A legjobb teljesítményt a  $H$ -teszt és/vagy a  $W$ -teszt nyújtotta, az  $F$ -teszt mindig gyengébb volt.

A kevésbé kockázatos időszakokban voltak ugyan olyan részvényhozamok, melyek jó közelítéssel normális eloszlásúak voltak, olyan eset azonban sohasem állt elő, hogy ez a portfólió minden részvényére egyidejűleg fennállt volna. Ez pedig azt jelenti, hogy az együttes normalitás hipotézise elméletileg nem megalapozott. A peremekre vonatkozó vizsgálat elég meggyőző annak



eldöntéséhez, hogy olyan modelleket itt nem alkalmazhatunk, melyek a több-dimenziós normalitás teljesülését megkövetelik.

## I. Függelék

Illusztrációképpen közlünk néhány számszerű eredményt. A bal oldali táblázatok az  $m = 10$   $n = 30$ , a jobb oldali táblázatok pedig az  $m = 10$   $n = 50$  esetet ábrázolják.

Az összes eset ( $m = 2 - 15$ ,  $n = 20, 30, 40, 50$ ) közlése helyhiány miatt nem lehetséges, de nem is lenne célszerű. Az alább kiválasztott két eset nagyjából jól reprezentálja a vizsgálat eredményét. A leginkább észrevehető eltérés — értelemszerűen — a kisebb minták ( $n = 30$ ) esetén van.

22. táblázat. Többváltozós  $\chi^2$  eloszlás  $\chi_3^2$  peremekkel

	10%	5%	1%		10%	5%	1%
$F$	0.9998	0.9996	0.9950	$F$	1.0002	1.0002	1.0002
$L$	0.9982	0.9968	0.9916	$L$	1.0002	1.0002	1.0002
$N$	0.9990	0.9976	0.9918	$N$	1.0002	1.0002	1.0002
$T$	0.9922	0.9768	0.8712	$T$	1.0002	1.0002	0.9992
$W$	0.7812	0.6702	0.4138	$W$	0.9478	0.9066	0.7542
$H$	1.0002	1.0002	1.0002	$H$	1.0002	1.0002	1.0002

23. táblázat. Többváltozós  $\chi^2$  eloszlás  $\chi_4^2$  peremekkel

	10%	5%	1%		10%	5%	1%
$F$	0.9980	0.9956	0.9844	$F$	1.0002	1.0002	1.0002
$L$	0.9926	0.9884	0.9692	$L$	1.0002	1.0002	1.0002
$N$	0.9958	0.9916	0.9712	$N$	1.0002	1.0002	1.0002
$T$	0.9778	0.9510	0.7890	$T$	1.0002	1.0000	0.9926
$W$	0.7490	0.6316	0.3788	$W$	0.9444	0.8894	0.7272
$H$	1.0002	1.0000	0.9982	$H$	1.0002	1.0002	1.0002

24. táblázat. Többváltozós  $\chi^2$  eloszlás  $\chi_5^2$  peremekkel

	10%	5%	1%		10%	5%	1%
$F$	0.9972	0.9932	0.9750	$F$	1.0002	1.0002	1.0002
$L$	0.9884	0.9816	0.9504	$L$	1.0002	1.0002	1.0002
$N$	0.9942	0.9864	0.9580	$N$	1.0002	1.0002	1.0002
$T$	0.9684	0.9310	0.7536	$T$	0.9998	0.9996	0.9876
$W$	0.7248	0.6058	0.3570	$W$	0.9292	0.8698	0.6872
$H$	0.9934	0.9882	0.9682	$H$	1.0002	1.0002	1.0002

25. táblázat. Többváltozós  $t$ -eloszlás,  $df = 3$

	10%	5%	1%		10%	5%	1%
$F$	0.7076	0.6272	0.4570	$F$	0.9588	0.9340	0.8678
$L$	0.5486	0.4658	0.3146	$L$	0.9100	0.8782	0.7922
$N$	0.5924	0.4832	0.3004	$N$	0.9240	0.8860	0.7756
$T$	0.6690	0.5792	0.4044	$T$	0.9204	0.8716	0.7544
$W$	0.9640	0.9342	0.8370	$W$	0.9982	0.9962	0.9804
$H$	0.9480	0.9260	0.8706	$H$	0.9950	0.9918	0.9814

26. táblázat. Többváltozós Cauchy-eloszlás

	10%	5%	1%		10%	5%	1%
$F$	0.9992	0.9976	0.9906	$F$	1.0002	1.0002	1.0002
$L$	0.9926	0.9876	0.9798	$L$	1.0002	1.0002	1.0000
$N$	0.9924	0.9888	0.9762	$N$	1.0002	1.0002	1.0002
$T$	0.9962	0.9920	0.9710	$T$	1.0002	1.0002	1.0000
$W$	1.0002	1.0000	0.9998	$W$	1.0002	1.0002	1.0002
$H$	1.0002	1.0002	1.0000	$H$	1.0002	1.0002	1.0002

27. táblázat. Többváltozós lognormális eloszlás

	10%	5%	1%		10%	5%	1%
$F$	1.0002	1.0002	1.0002	$F$	1.0002	1.0002	1.0002
$L$	1.0002	1.0002	1.0002	$L$	1.0002	1.0002	1.0002
$N$	1.0002	1.0002	1.0002	$N$	1.0002	1.0002	1.0002
$T$	1.0002	1.0002	1.0000	$T$	1.0002	1.0002	1.0002
$W$	0.9914	0.9816	0.9316	$W$	1.0002	1.0002	0.9994
$H$	1.0002	1.0002	1.0002	$H$	1.0002	1.0002	1.0002

28. táblázat. Többváltozós kevert normális eloszlás (0.8, 9.0)

	10%	5%	1%		10%	5%	1%
$F$	0.7866	0.7102	0.5342	$F$	0.9864	0.9774	0.9400
$L$	0.6232	0.5440	0.3742	$L$	0.9612	0.9414	0.8812
$N$	0.6656	0.5672	0.3662	$N$	0.9656	0.9446	0.8706
$T$	0.7534	0.6548	0.4238	$T$	0.9694	0.9446	0.8216
$W$	0.9946	0.9912	0.9418	$W$	1.0002	1.0002	0.9990
$H$	0.9938	0.9892	0.9734	$H$	0.9994	0.9994	0.9990

29. táblázat. Többváltozós kevert normális eloszlás (0.9,16.0)

	10%	5%	1%
<i>F</i>	0.7826	0.7184	0.5822
<i>L</i>	0.6266	0.5614	0.4228
<i>N</i>	0.6338	0.5420	0.3688
<i>T</i>	0.7936	0.7334	0.5878
<i>W</i>	0.9604	0.9580	0.9498
<i>H</i>	0.9590	0.9568	0.9488

	10%	5%	1%
<i>F</i>	0.9694	0.9548	0.9220
<i>L</i>	0.9336	0.9154	0.8638
<i>N</i>	0.9310	0.8978	0.8226
<i>T</i>	0.9640	0.9526	0.9068
<i>W</i>	0.9954	0.9952	0.9940
<i>H</i>	0.9936	0.9934	0.9914

30. táblázat. Többváltozós Burr-Pareto-Logisztikus eloszlás

	10%	5%	1%
<i>F</i>	0.9904	0.9786	0.9286
<i>L</i>	0.9704	0.9470	0.8844
<i>N</i>	0.9818	0.9648	0.9032
<i>T</i>	0.9236	0.8270	0.5000
<i>W</i>	0.2206	0.1318	0.0382
<i>H</i>	1.0002	1.0002	1.0002

	10%	5%	1%
<i>F</i>	1.0002	1.0002	1.0000
<i>L</i>	1.0002	1.0002	0.9996
<i>N</i>	1.0002	1.0002	0.9996
<i>T</i>	0.9990	0.9944	0.9186
<i>W</i>	0.1824	0.0974	0.0216
<i>H</i>	1.0002	1.0002	1.0002

## Irodalomjegyzék

1. Anderson, T. W. [1958]: Multivariate Statistical Analysis; Wiley, New York.
2. Andrews, D.F., Gnanadesikan, R. és Warner, J.L. [1973]: Methods for Assessing Multivariate Normality; In: Krishnaiah, P.R. (ed.): Multivariate Analysis III., Academic Press, New York.
3. Artzner, P.-Delbaen, F, Eber, M. és Heath, D. [1999]: Coherent Measures of Risk; *Mathematical Finance*, **9**, pp.203-208.
4. Baringhaus, L. és Henze, N. [1991]: Limit Distributions for the Measures of Multivariate Skewness and Kurtosis Based on Projections; *Journal of Multivariate Analysis*, **38**, pp.51-69.
5. Baringhaus, L. és Henze, N. [1992]: Limit Distributions for Mardia's Measure of Multivariate Skewness; *Annals of Statistics*, **20**, pp.1889-1902.
6. Benedek, G.[1999]: Opcióárazás numerikus módszerekkel, *Közgazdasági Szemle*, **46**, pp.905-930.
7. Benedek, G.-Kóbor, Á. és Pataki, A. [2001]: Modeling Dependence with  $m$ -Variate Copulas and Applications to Equity Portfolios, *Kézirat*.
8. Bera, A.K. és Jarque, C.M. [1981]: An Efficient Large Sample Test for Normality of Observation and Regression Residuals; *Working Papers in Economics* **40**, Australian National University.
9. Bera, A.K. és John, S. [1983]: Tests for Multivariate Normality with Pearson Alternatives; *Communications in Statistics - Theory and Methods*, **12**, pp.103-117.
10. Bickel, P.J. és Doksum, K.A. [1977]: Mathematical Statistics: Basic Ideas and Selected Topics; Holden Day, San Francisco.
11. Black, F. és Scholes, M. [1973]: The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**, pp. 637-654.
12. Churchill, R.V. [1963]: Fourier Series and Boundary Value Problems, 2. kiadás. McGraw-Hill, New York.
13. Conn, A.R., Gould, N.I.M., és Toint, Ph.L. [1992]: LANCELOT: A Fortran Package for Large-Scale Nonlinear Optimization (Release A); Springer-Verlag, Berlin Heidelberg.

14. Cox, J. és Ross, S. [1976]: The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, **3**. március, pp. 145-166.
15. Cox, J.-Rubinstein, M. és Ross, S. [1979]: Option Pricing A Simplified Approach; *Journal of Financial Economics*, **7**, augusztus, pp. 229-263.
16. Davies, C.S. és Stephens, M.A. [1978]: Algorithm AS128. Approximating the Covariance Matrix of Normal Order Statistics; *Applied Statistics*, **27**, pp.206-212.
17. Embrechts, P.-McNeil, A. és Strauman, D. [1999]: Correlation and Dependence in Risk Management: Properties and Pitfalls; Preprint ETH, Zürich.
18. Fang, K.T.-Kotz, S. és Ng, K.W. [1990]: Symmetric Multivariate and Related Distributions; Chapman and Hall, London.
19. Filliben, J.J. [1975]: The Probability Plot Correlation Coefficient Test for Normality; *Technometrics*, **17**, pp.111-117.
20. Gablonsky, J.M. [2001]: DIRECT 2.0 User Guide. North Carolina State University, Dept. of Mathematics.
21. Galambos, J. [1975]: Order Statistics of Samples from Multivariate Distributions; *Journal of the American Statistical Association*, **70**, pp. 674-680.
22. Gnanadesikan, R. [1977]: Methods for Statistical Data Analysis for Multivariate Observations, New York, Wiley.
23. Gnanadesikan, R. és Kettenring, J.R. [1972]: Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data; *Biometrika*, **28**, pp.81-124.
24. Gumbel, E.J. [1960]: Distributions des valeurs extremes en plusieurs dimensions. *Publications of the Institute of Statistics, University of Paris*, **9**, pp. 171-173.
25. Healy, M.J.R. [1968]: Multivariate Normal Plotting; *Applied Statistics* **17**, pp. 157-161.
26. Hull, J.C. [1993]: Options, Futures, and other Derivative Securities. 2. kiadás. Prentice-Hall International, Inc., Engle Wood Cliffs, New Jersey.
27. Hull, J.C. és White, A. [1987]: The Pricing of Options on Assets with Stochastic Volatilities; *The Journal of Finance*, június, pp. 281-300.

28. Jeffreys, H. [1939]: Theory of Probability; Oxford: Clarendon Press.
29. Joe, H. [1997]: Multivariate Models and Dependence Concepts; Chapman and Hall, London.
30. Kendall, M.G. és Gibbons, J.D. [1990]: Rank Correlation Methods; 5th Edition, Griffin, London.
31. Knuth, D.E. [1981]: The Art of Computer Programming: Seminumerical Algorithms; Reading, MA, Addison-Wesley.
32. Kupiec, P. [1995]: Techniques for Verifying the Accuracy of Risk Measurement Models; *The Journal of Derivatives*, **3**, pp. 73-84.
33. Lehmann, E.L. [1966]: Some concepts of Dependence; *Annals of Mathematical Statistics* **37**, pp. 1137-1153.
34. Lindskog, F. [2000a]: Linear Correlation Estimation; *RiskLab Report*, ETH-Zentrum, Zürich.
35. Lindskog, F. [2000b]: Modeling Dependence with Copulas and Applications to Risk Management; Master Thesis, ETH-Zentrum, Zürich.
36. Mahibbur Rahman, M.-Govindarajulu, Z.[1997]: A Modification of the Test of Shapiro and Wilk for Normality; *Journal of Applied Statistics*, **24**, pp.219-235.
37. Malkovich, J.F.-Afifi, A.A.[1973]: On Tests for Multivariate Normality; *Journal of the American Statistical Association*, **68**, pp.176-179.
38. Mardia, K.V. [1970]: Measures of Multivariate Skewness and Kurtosis with Applications; *Biometrika*, **57**, pp.519-530.
39. Mardia, K.V. [1980]: Test of Univariate and Multivariate Normality; In: Krishnaiah, P.R. (ed.): Handbook of Statistics, Vol. 1., Analysis of Variance, Elsevier Science Publishers, B.V.
40. Mardia, K.V. és Foster, K. [1983]: Omnibus Tests of Multivariate Normality Based on Skewness and Kurtosis; *Communications in Statistics - Theory and Methods*, **12**, pp.207-221.
41. Markowitz, H. [1959]: Portfolio Selection; *The Journal of Finance*, **18**, March.
42. McKay, M.D.- Beckman, R.J. és Conover, W.J. [1979]: A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code; *Technometrics*, **21**, pp. 239-245.

43. Móry F.T. és Székely, G. (szerk.) [1986]: Többváltozós statisztikai analízis. Műszaki Könyvkiadó, Budapest.
44. Mudholkar, G.S. - Srivasta, D.K. és Lin, T. [1995]: Some p-variate Adaptations of the Shapiro-Wilk Test of Normality, *Communicating Statistics*, **24**, pp. 953-985.
45. Nelson, B.L. [1990]: Control Variate Remedies; *Operations Research*, **38**, pp.974-992.
46. Owen, A.B. [1992a]: A Central Limit Theorem for Latin Hypercube Sampling; *Journal of the Royal Statistical Society, Series B*, **54**, pp. 514-551.
47. Owen, A.B. [1992b]: Orthogonal Arrays for Computer Integration and Visualization; *Statistica Sinica*, **2**, pp. 2-2.
48. Palágyi, Z. [2001]: Piaci hozamok modellezése stabil eloszlásokkal; Ph.D. Értekezés, Budapesti Közgazdasági és Államigazgatási Egyetem.
49. Pataki, A. [2001]: A többváltozós Shapiro-Wilk tesztek vizsgálata; *Sigma* (elfogadva)
50. Pearson, E.S., D'Agostino, R.B. és Bowman, K.O. [1977]: Tests for Departure from Normality: Comparison of Powers; *Biometrika*, **64**, pp.231-246.
51. Pintér, J. [1996]: Global Optimization in Action: Continuous and Lipschitz Optimization: Algorithms, Implementation and Applications; Kluwer Academic Publisher.
52. Press, W.H.-Teukolsky, S.A.-Vetterling, W.T.-Flannery, B.P. [1992]: Numerical Recipes in C: The Art of Scientific Computing (2nd ed.); Cambridge University Press, New York
53. Roy, S.N. [1953]: On a Heuristic Method of the Test Construction and Its Use in Multivariate Analysis; *Annals of Mathematical Statistics*, **24**, pp.220-238.
54. Roy, S.N. [1957]: Some Aspects of Multivariate Normality; Wiley, New York.
55. Royston, P.[1982a]: An Extension of Shapiro and Wilk's Test for Normality to Large Samples; *Applied Statistics*, **31**, pp.115-124.
56. Royston, P.[1982b]: Algorithm AS181: The W-test for Normality; *Applied Statistics*, **31**, pp.176-180.



57. Royston, P.[1982c]: Algorithm AS177: Expected Normal Order Statistics (exact and approximate); *Applied Statistics*, **31**, pp.161-175.
58. Royston, P. [1983]: Some techniques for assessing multivariate normality based on Shapiro-Wilk W-test; *Applied Statistics*, **32**, pp.121-133.
59. Royston, P.[1992]: Approximating the Shapiro-Wilk W-test for non-normality; *Statistics and Computing*, **2**, pp.117-119.
60. Royston, P.[1995]: A Remark on Algorithm AS181: The W-test for Normality; *Applied Statistics*, **44**, pp.547-551.
61. Royston, P.-Altman, D.G.[1994]: Regression using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling; *Applied Statistics*, **43**, pp.429-467.
62. Royston, P.-Altman, D.G.[1997]: Approximating Statistical Functions by using Fractional Polynomial Regression; *The Statistician*, **46**, pp.411-422.
63. Sarhan, A.E. és Greenberg, B.G. [1956]: Estimation of Location and Scale Parameters by Order Statistics from Single and Double Censored Samples; *Annals of Mathematical Statistics*, **27**, pp.427-451.
64. Shapiro, S.S.-Wilk, M.B.[1965]: An Analysis of Variance Test for Normality (complete samples); *Biometrika*, **52**, pp.591-611.
65. Shapiro, S.S.-Wilk, M.B., és Chen, H.J. [1968]: Comparative Study of Various Tests for Normality; *Journal of the American Statistical Association*, **63**, pp.1343-1372.
66. Shea, B.L.és Scallan, A.J. [1988]: AS R72. A Remark on Algorithm AS128: Approximating the Covariance Matrix of Normal Order Statistics; *Applied Statistics*, **37**, pp.151-155.
67. Sklar, A. [1996]: Random Variables, Distribution Functions and Copulas — a personal look backward and forward; In: Rüschendorf, L-Schweitzer, B. és Taylor, M.D.: Distributions with Fixed Marginals and Related Topics, Hayward CA., Institute of Mathematical statistics, pp.1-14.
68. Stein, M. [1987]: Large Sample Properties of Simulations Using Latin Hypercube Sampling; *Technometrics* **29**, pp. 143-151.
69. Tew, J.D. és Wilson, J.R [1992]: Validation of Simulation Analysis Methods for the Schruben-Margolin Correlation Induction Strategy; *Operations Research*, **40**, pp.87-103.

70. Wichura, M.J. [1988]: Algorithm AS241: The percentage points of the normal distribution; *Applied Statistics*, **37**, pp.477-484.
71. Wolff, R.W. [1989]: Stochastic Modeling and the Theory of Queues; Prentice Hall, Englewood Cliffs, N.J.
72. Zhigljavsky, A.A. [1991]: Theory of Global Random Search; (*szerk.* Pintér János), Kluwer Academic Publisher.